Humboldt-Universität zu Berlin - Geographisches Institut

CHANGE DETECTION FOR APPLICATION IN URBAN GEOGRAPHY BASED ON VERY HIGH RESOLUTION REMOTE SENSING

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

im Fach Geographie

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

Tobias Leichtle (M.Sc.)

Präsidentin der Humboldt-Universität zu Berlin

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät

Prof. Dr. Elmar Kulke

Gutachter/innen: 1. Prof. Dr. Tobia Lakes

2. Prof. Dr. Birgit Kleinschmit

3. Prof. Dr. Andreas Schmitt

Tag der mündlichen Prüfung: 19.12.2019

Acknowledgements

At this point, I would like to express my sincere gratitude to everyone who was involved and contributed to this thesis.

First and foremost, I would like to thank my supervisors, Prof. Dr. Tobia Lakes and Dr. Hannes Taubenböck for their continuous and encouraging support as well as their very valuable scientific advice during all stages of this dissertation and for making this work possible.

My very special thanks go to the Company for Remote Sensing and Environmental Research (SLU) for funding large parts of this dissertation through different projects and its owner, Dr. Klaus Martin, for his continuous support as well as the possibility and the freedom of conducting research besides work.

Special thanks go to the German Remote Sensing Data Center of the German Aerospace Center, in particular to the department of Geo-Risks and Civil Security and its head Prof. Günter Strunz for the opportunity to create my thesis in an exciting scientific environment.

The basis for this thesis was established in the BMBF-funded research project DELIGHT (Delta Information System for geoenvironmental and human Habitat Transition, www.delight.eoc.dlr.de), which provided inspiration and ideas for conducting this work. In addition, very high resolution (VHR) remote sensing imagery and field data from DELIGHT was used in this thesis.

I also want to thank European Space Imaging (EUSI, www.euspaceimaging.com) for providing additional VHR images for my work.

Source of inspiration and motivation in everyday work as well as numerous opportunities for discussion came from all my colleagues of the team City and Society within the department of Geo-Risks and Civil Security, particular thanks go to Dr. Christian Geiß and Dr. Michael Wurm for valuable advice and critical discussions on my work and possible pathways towards a PhD.

I am particular grateful to my family for encouragement and uncontested support.

Finally, my deep gratitude goes to my wife Sahra Abdullahi for her endless patience, scientific advice, continuous support, and Nora.

Abstract

The continuous and ongoing transformation of the Earth's surface can be observed especially in urban areas, which are hot spots of global change. In addition to the process of urbanization, the world population has increased exponentially since 1900, while future population growth is entirely projected to cities. Facing these trends, the envisaged goal is to promote efficient, resilient, and sustainable development of cities through politics and urban planning. In this regard, suitable highly detailed and up-to-date information is required, which can be delineated based on various earth observation sensors.

In this context, this thesis aims at the development of a change detection approach for monitoring urban environments based on optical remote sensing data with very high resolution (VHR) and consequent exemplary application of the assessment of the ghost city phenomenon in the context of urban geography. The unsupervised object-based change detection approach captures the construction of individual buildings and returned viable results in the order of 0.8 to 0.9 according to kappa statistics in the city of Dongying, China. The methodology utilizes object-based difference features based on existing building geometries and employs clustering as the comparison method for the delimitation of changed and unchanged buildings. Several different clustering algorithms were evaluated and their ability to separate possible arbitrary distribution of clusters in the feature space was documented. The proposed approach mitigates differences in the acquisition system and is capable of handling VHR data from different sensors, with deviating viewing geometries, in different acquisition modes, and/or in different seasons. This allows the utilization of all present and future available sources of VHR data for monitoring urban areas at medium and small spatial scale. The transferability of the approach is investigated with particular focus on the nature and effects of class distribution and potentially occurring class imbalance in this thesis. For this purpose, a diagnostic framework is developed and consequently applied in the city of Dongying, China and the city of Munich, Germany, which are characterized by different built-up structure as well as divergent temporal evolution. The results showed that situations of imbalanced class distribution generally provide less reliable identification of changes compared to balanced or close to balanced situations. However, clustering techniques like the genetic k-means algorithm, which account for non-linearity and possess capabilities for the detection of clusters of arbitrary shape, provide more robust results compared to the traditional k-means algorithm. In addition, the diagnostic framework can be utilized for further development and

application of present and future methods for change detection and image classification in order to assess and improve robustness and transferability of the methodology. The assessment of the presence or absence of the ghost city phenomenon is conducted as an exemplary application of urban geography in the city of Dongying, China. The conceptual framework based on VHR imagery and census counts in this thesis replicates undercapacity with respect to the residential population as one of the key characteristics of a ghost city. A 4d functional city model is established and serves for estimation of available living space and population capacity of residential buildings. Subsequently, population capacity estimates are related to actual permanent residential population from census counts revealing a significant mismatch and thus, high likelihood for the emergence and presence of the ghost city phenomenon within the urban area of Dongying. In addition, the proposed approach allows identification of specific regions of the urban area that are most likely affected. In perspective, the 4d functional city model also enables other fields of application in the context of changing urban environments.

Accordingly, this thesis not only discusses the applicability of heterogeneous VHR data sets for unsupervised change detection in urban areas as well as the transferability of the methodology regarding the distribution of classes, but also contributes to a more detailed understanding of urban growth in highly dynamic Chinese cities based on the exemplary application of the ghost city phenomenon in the city of Dongying.

Zusammenfassung

Städte sind Brennpunkte des globalen Wandels. Seit etwa 1900 steigt die Weltbevölkerung exponentiell an, wobei das zukünftig prognostizierte Bevölkerungswachstum vollständig auf Städte entfallen wird. Vor den damit verbundenen Herausforderungen ist es das angestrebte Ziel, eine effiziente, widerstandsfähige und nachhaltige Stadtentwicklung durch Politik und Stadtplanung zu fördern. In diesem Zusammenhang sind hochdetaillierte und aktuelle Informationen über die Entwicklung von Städten zwingend erforderlich. Hierfür liefern moderne Erdbeobachtungssensoren eine ideale Datenbasis zum Monitoring von Städten.

Vor diesem Hintergrund wird im Rahmen der vorliegenden Arbeit ein Verfahren zur Änderungserkennung auf Basis höchstaufgelöster optischer Erdbeobachtungsdaten entwickelt und anschließend im stadtgeographischen Anwendungsfall zur Bewertung einer potenziell vorliegenden Geisterstadt in Wert gesetzt. Das unüberwachte objektbasierte Verfahren erfasst die Entstehung neuer Gebäude mit einer Genauigkeit von 0,8 bis 0,9 entsprechend der Kappa Statistik in einem Testgebiet in der chinesischen Stadt Dongying. Dabei werden objektbasierte Differenzmerkmale auf Basis vorhandener Gebäudegeometrien eingesetzt und mittels Clustering veränderte und unveränderte Gebäude unterschieden. Verschiedene Clustering-Algorithmen werden analysiert, um komplexe Klassenverteilungen im Merkmalsraum trennen zu können. Ein Vorteil des entwickelten Ansatzes ist die Nutzung höchstaufgelöster Aufnahmen verschiedener Sensoren mit unterschiedlichen Aufnahmegeometrien, verschiedenen Erfassungsmodi und/oder zu unterschiedlichen Jahreszeiten. Dies ermöglicht die Nutzung des gesamten Datenbestandes aktueller und zukünftig verfügbarer höchstaufgelöster Satellitenbilddaten zum Monitoring von Städten auf mittleren und kleinen räumlichen Skalen. Die Übertragbarkeit des Ansatzes zur Änderungserkennung wird mit besonderem Augenmerk auf die Verteilung und das mögliche Ungleichgewicht der Klassen untersucht. Zu diesem Zweck wird ein Rahmenwerk entwickelt und in der chinesischen Stadt Dongying sowie der deutschen Stadt München angewandt, welche sich durch unterschiedliche Bebauungsstruktur sowie eine unterschiedliche zeitliche Dynamik auszeichnen. Die Ergebnisse zeigen geringere Genauigkeiten bei ungleich verteilten Klassen im Gegensatz zu einer ausgewogenen Verteilung. Clustering-Techniken wie der genetic k-means Algorithmus, die nichtlineare Klassengrenzen berücksichtigen und Cluster beliebiger Form erfassen, lieferten robustere Ergebnisse im Vergleich zum traditionellen k-means Algorithmus. Darüber hinaus kann das Rahmenwerk für die Weiterentwicklung aktueller und zukünftiger Methoden zur

Änderungserkennung und Klassifikation genutzt werden, um die Robustheit und Übertragbarkeit der Methodik zu beurteilen und zu verbessern. Die Bewertung potenziell vorliegender Geisterstädte wird als exemplarische stadtgeographische Anwendung am Beispiel der chinesischen Stadt Dongying gezeigt. Das Bewertungskonzept basiert auf der Annahme, dass eine geringe Auslastung des verfügbaren Wohnraums eines der wichtigsten Merkmale einer Geisterstadt darstellt und auf Grundlage höchstaufgelöster Satellitenbilddaten in Kombination mit Zensusinformation analysiert werden kann. Dazu wird ein funktionales 4D-Stadtmodell erstellt, welches zur Abschätzung der verfügbaren Wohnfläche und damit der potenziellen Einwohnerzahl von Wohngebäuden dient. Anschließend werden die Schätzungen der Bevölkerungskapazität mit der tatsächlichen permanenten Wohnbevölkerung der Zensusdaten verglichen. Aufgrund signifikanter Unterschiede dieser beiden Größen ergibt sich eine hohe Wahrscheinlichkeit für die Entstehung einer Geisterstadt in der Stadt Dongying. Darüber hinaus ermöglicht der Ansatz die Identifikation bestimmter Regionen des Stadtgebiets, die am wahrscheinlichsten betroffen sind. Perspektivisch lässt sich das funktionale 4D-Stadtmodell auch für andere Anwendungsbereiche im Kontext sich verändernder urbaner Räume einsetzen.

Die vorliegende Arbeit diskutiert nicht nur die Anwendbarkeit heterogener, höchstaufgelöster Satellitenbilddaten für die unüberwachte Änderungserkennung in Städten sowie die Übertragbarkeit der Methodik bezüglich der Verteilung der Klassen, sondern trägt darüber hinaus auf Basis der Bewertung potenziell vorliegender Geisterstädte am Beispiel der Stadt Dongying zu einem detaillierteren Verständnis des Stadtwachstums in hochdynamischen chinesischen Städten bei.

Contents

List of Figures

List of Tables

Abbreviations

ATCOR	Atmospheric and Topographic Correction
CVA	change vector analysis
DELIGHT	Delta Information System for geoenvironmental and human Habitat Transition
DLR	German Aerospace Center
DMSP/OLS	Defense Meteorological Satellite Program / Operational Linescan System
DSM	digital surface model
DTM	digital terrain model
EUSI	European Space Imaging
EO	earth observation
EM	expectation–maximization
GHSL	Global Human Settlement Layer
GLCM	gray-level co-occurrence matrix
GLDV	gray-level difference vector
GUF	Global Urban Footprint
GPW	Gridded Population of the World
GRUMP	Global Rural-Urban Mapping Project
HR	high resolution
IPCC	Intergovernmental Panel on Climate Change
IPBES	Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services
LBS	location based services
MOD500	500 m Map of Global Urban Extent
MERIS	MEdium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
MS	multispectral
nDSM	normalized digital surface model
NIR	near infrared

NDVI	normalized difference vegetation index
OA	overall accuracy
OBIA	object-based image analysis
OSM	OpenStreetMap
PAM	partitioning around medoids
PR	precision-recall
PC	principal component
PCA	principal component analysis
PPV	positive predictive value
RF	random forest
RMSE	root mean square error
ROC	receiver operator characteristics
SAR	synthetic aperture radar
SAVI	normalized difference soil index
SLU	Company for Remote Sensing and Environmental Research
SMOTE	synthetic minority oversampling technique
SOM	self-organizing map
SVM	support vector machines
TPR	true positive rate
TNR	true negative rate
TSS	true-skill-statistics
UN-Habitat	United Nations Human Settlements Programme
VHR	very high resolution
WBGU	German Advisory Council on Global Change
YRD	Yellow River Delta

Chapter I. Introduction

1. Global transformation of cities

The Earth's surface is under continuous and ongoing transformation. Especially in recent times the rate of change has been accelerating, whereas it is generally acknowledged that most of these alterations can be accounted to humanity (Vitousek et al. 1997). In this context of global change, the tremendous anthropogenic influences on ecosystems worldwide led to the introduction of the 'Anthropocene' as a new era (Crutzen 2002). Although the 'Anthropocene' is not yet officially recognized as an unit of geological time (Lewis and Maslin 2015), global human activities have become pervasive and profound at unprecedented magnitudes and rates, especially after the end of the Second World War (Steffen et al. 2007). The clear influence of humanity on this rapid global transformation was also ascertained by the German Advisory Council on Global Change (WBGU) (WBGU 2016), the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) (IPBES 2019), as well as the Intergovernmental Panel on Climate Change (IPCC), in particular with respect to climatic changes on Earth (IPCC 2014).

These processes of reshaping and modification of the Earth's surface can be observed especially in cities and their surroundings, which are hot spots and drivers of environmental changes in many regards (Grimm et al. 2008). Within the past 100 years, global population has increased exponentially and the future global population growth is entirely predicted for cities (United Nations 2017). With regard to urban areas, in 1900 only 15% of the world's population lived in cities, whereas today over 50% of the people live in urban environments (Spence et al. 2009). In general, population growth is predicted for urban areas of all sizes, i.e., from small settlements smaller than 500,000 dwellers to megacities and megaregions with 10 million inhabitants or more (United Nations 2014). However, urban growth is not equally distributed across the globe as most of the current urban as well as total population is concentrated in Asia, whereas the greatest increase in population is expected for Africa (United Nations 2017). In 2100, Asia and Africa will be home of around 4.5 billion people each, compared to an estimated total global population of 11.2 billion (United Nations 2017).

A well-established program in the context of urbanization and population growth is the United Nations Human Settlements Programme (UN-Habitat, www.unhabitat.org), which was launched in the 1970s aiming at

the elaboration of strategies for efficient and sustainable cities against the background of rapid urban development. Another institutional approach for facing these trends are the Sustainable Development Goals introduced by the United Nations (www.un.org/sustainabledevelopment), which include a specific objective with respect to urban development (i.e., Goal 11 "Sustainable Cities and Communities") for promotion of inclusive, safe, resilient, and sustainable cities. Finally, the World Bank (www.worldbank.org) collects data on urban development worldwide and provides knowledge and recommendations for governments through its regularly updated reports on regional as well as country-specific urbanization reviews.

In order to promote the envisaged goal of efficient, resilient, and sustainable development of cities, appropriate information is urgently required for political decision makers and urban planning. In addition to other data sources like official statistics, in situ measurements, or crowdsourcing data, suitable and relevant (geo-)information is available through data collected by various sensors (Scholten and Stillwell 2013). While data and corresponding information are relatively well available for cities of the Western world, there is an increased need for appropriate information especially in emerging and developing countries of Asia and Africa, as these areas are hot spots of urbanization and projected population growth (Cohen 2006; UN-Habitat 2016; World Bank 2013).

2. Remote sensing for monitoring of urban areas

Against the background of global environmental changes and rapid urban development, remote sensing has proven an adequate tool for observation and monitoring (i.e., regularly repeated observations of an area over a period of time) of the Earth's surface in many regards (Bartholomé and Belward 2005; Buiten and Clevers 1994; Townshend et al. 1991), also with respect to urban areas (Taubenböck et al. 2015; Weng et al. 2018; Yang 2011). Based on numerous research studies and continuous improvement of methods for information extraction in urban environments, several global urban products and applications at medium and small spatial scales have been developed using remote sensing imagery within the past decades.

Examples of global products for the assessment of the urban extent with moderate spatial resolution are the 500 m Map of Global Urban Extent (MOD500) based on MODIS (Moderate Resolution Imaging Spectroradiometer) data (Schneider et al. 2009) or the 300 m GlobCover 2009 which was delineated from MERIS (MEdium Resolution Imaging Spectrometer) imagery (Arino et al. 2007; Bontemps et al. 2010). Two recent high resolution (HR) global urban maps are the 12 m Global Urban Footprint (GUF) product which was

produced from TerraSAR-X and TanDEM-X SAR (Synthetic Aperture Radar) data (Esch et al. 2013) and the 10 m Global Human Settlement Layer (GHSL) based on high resolution optical imagery from different sensors (Pesaresi et al. 2013). Other relevant data sources and approaches for mapping urban areas at medium and small spatial scales include multi-temporal Landsat imagery (Taubenböck et al. 2012) or the DMSP/OLS (Defense Meteorological Satellite Program / Operational Linescan System) nighttime light data (Zhang and Seto 2011). In addition, the Sentinel-1 (C-band SAR sensor) and Sentinel-2 (multispectral sensor) missions have been providing remote sensing data with high spatial resolution for global urban mapping recently (Ban et al. 2017). These remote sensing data have also been used in combination with census counts for population estimation and mapping to provide global products like GPW (Gridded Population of the World) (Doxsey-Whitfield et al. 2015; Tobler et al. 1997), GRUMP (Global Rural-Urban Mapping Project) (Balk et al. 2005), LandScan (Dobson et al. 2000), or the WorldPop data set (Stevens et al. 2015) at medium and high spatial resolution.

Satellite remote sensing imagery for multi-temporal assessment of urban areas is available with large spatial coverage from several space-borne sensors since the 1970s (Townshend et al. 1991). These sensors possess diverse characteristics with respect to temporal (i.e., imaging frequency, revisit time), spectral (i.e., number and wavelength of spectral bands, radiometric resolution, and radiometric calibration) as well as spatial (i.e., spatial resolution, and geometric registration) properties of the acquired data (Townshend and Justice 1988). In this regard, comprehensive capabilities have been demonstrated in literature for urban monitoring at large and global scales as mentioned above. Correspondingly, several effective methods for automatic and unsupervised change detection (i.e., identification of the occurrence and/or the type of changes within a defined time frame) based on satellite imagery with moderate (i.e., several 100 m) and high (i.e., 10 - 100 m) spatial resolution have been proposed in literature (Coppin et al. 2004; Lu et al. 2004; Radke et al. 2005; Singh 1989).

Around the year 2000, a new generation of optical remote sensing satellites with very-high resolution (VHR), which are commonly defined according to spatial resolution of 1 m or less, became available. These data enable detailed characterization and analysis of the Earth's surface with particular advantages for mapping urban areas where various surface materials and objects are concentrated in a highly dynamic and complex manner (Carleer and Wolff 2006). In this context, VHR imagery allows the identification of individual man-made and natural objects and facilitates the classification of the intra-urban morphologic configuration of cities (Bellens et al. 2008; Wurm et al. 2011). After around 15 years of data collection by sensors like IKONOS, QuickBird, GeoEye-1/2, WorldView-1/2/3/4 or Pléiades-1/2, VHR imagery became more available especially for applications in the temporal domain. However, existing research studies based on multi-temporal VHR data mostly analyze only small spatial coverages (i.e., single scenes) and corresponding products with greater spatial coverage are widely

absent. In addition, the practical accessibility of this imagery remains challenging due to high cost of such commercial data at present. Nevertheless, the future increasing availability of VHR imagery, potentially decreasing data cost, as well as new VHR missions with alternative data policies (e.g., nano- and microsatellites like Planet SkySat) enable to value VHR data for information extraction and monitoring of urban areas (Butler 2014; Doncaster et al. 2016; Shao et al. 2017). Furthermore, the majority of research studies focus on change detection based on homogeneous data from only one sensor, whereas the combination of imagery from different sensors would significantly expand the potential VHR data basis. Thus, automatic and unsupervised change detection techniques, with additional capabilities of handling data from different sensors, are strongly required for effective analysis of the dynamics of intra-urban morphologic structure and large spatial coverage of multi-temporal VHR data (Bruzzone and Bovolo 2013).

2.1 Conceptualization of change in VHR remote sensing images

For the analysis and identification of changes in multi-temporal VHR imagery, several approaches and their applicability have been summarized according to different concepts of image analysis (e.g., Bhagat (2012), Hussain et al. (2013), Lu et al. (2014a), Tewkesbury et al. (2015)) in order to counter the complexity and heterogeneity of the image content in VHR data. To address the challenges caused by high dynamics in urban areas, Bruzzone and Bovolo (2013) proposed a versatile framework for change detection in VHR remote sensing imagery by introducing a taxonomy of possible radiometric changes which enables the association of their causes and semantic meaning. Based on this systematization, relevant changes to the specific application can be delimited against non-relevant radiometric differences and the methodology for change detection can be designed accordingly.

At the most generalized level of radiometric changes (Ω_{Rad}) that occur in between a pair of multi-temporal remote sensing images, changes due to acquisition conditions (Ω_{Acq}) can be separated from changes occurred on the ground (Ω_{Grd}) (Fig. I-1).

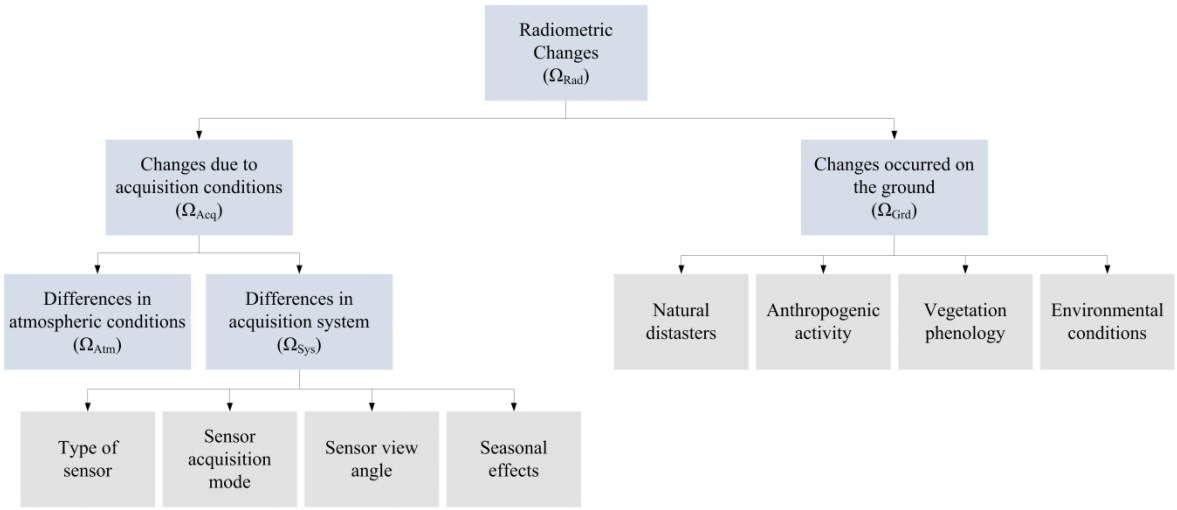

Fig. I-1: Conceptualization of change in multi-temporal VHR remote sensing images according to Bruzzone and Bovolo (2013).

According to the left side of Fig. I-1, changes due to acquisition conditions (Ω_{Acq}) are composed of differences in atmospheric conditions (Ω_{Atm}) as well as differences in the acquisition system (Ω_{Sys}). Varying atmospheric conditions mostly affect passive remote sensing imagery and include changing weather conditions (e.g., presence of clouds) as well as variations in atmospheric transmittance (e.g., concentration of ozone, water vapor, carbon dioxide, etc.). Differences in the acquisition system comprise different types of sensors, varying sensor view angles, seasonal effects (e.g., incidence angle of solar radiation), and deviating sensor acquisition modes of the VHR image pair. While atmospheric conditions can be mitigated by means of atmospheric correction as well as detection and masking of clouds, differences in the acquisition system are highly likely in a multi-temporal data setting of VHR imagery since suitable image pairs from the same sensor of a particular area of interest for a specific period of time with similar acquisition parameters are largely absent due to the intermittent acquisition tasking of most VHR remote sensing satellites. Especially different sensor view angles and changing solar incidence angles within the course of the year strongly affect the multi-temporal representation of objects which possess vertical extension in VHR imagery in terms of parallax distortion as well as object shadows.

The right side of Fig. I-1 depicts the changes that occurred on the ground (Ω_{Grd}), which are commonly relevant from the application point of view. These changes of land cover are categorized into four groups according to Bruzzone and Bovolo (2013), however, numerous other categorizations may be utilized dependent on the specific conceptualization and application context of change detection. In Fig. I-1, changes due to natural disasters (e.g., floods, earthquakes, landslides, etc.), vegetation phenology (e.g., crop growth, fall of leaves, etc.),

environmental conditions (e.g., desertification, glacier retreat, snow cover, etc.), and anthropogenic activity (e.g., forest clear cut, construction of buildings, etc.) are considered. Especially the latter category is highly relevant from the perspective of urban geography and urban growth. In particular, VHR images enable the detection of the built environment and its changes, since man-made objects on the land surface, such as individual buildings, are less recognizable in images with moderate or high spatial resolution.

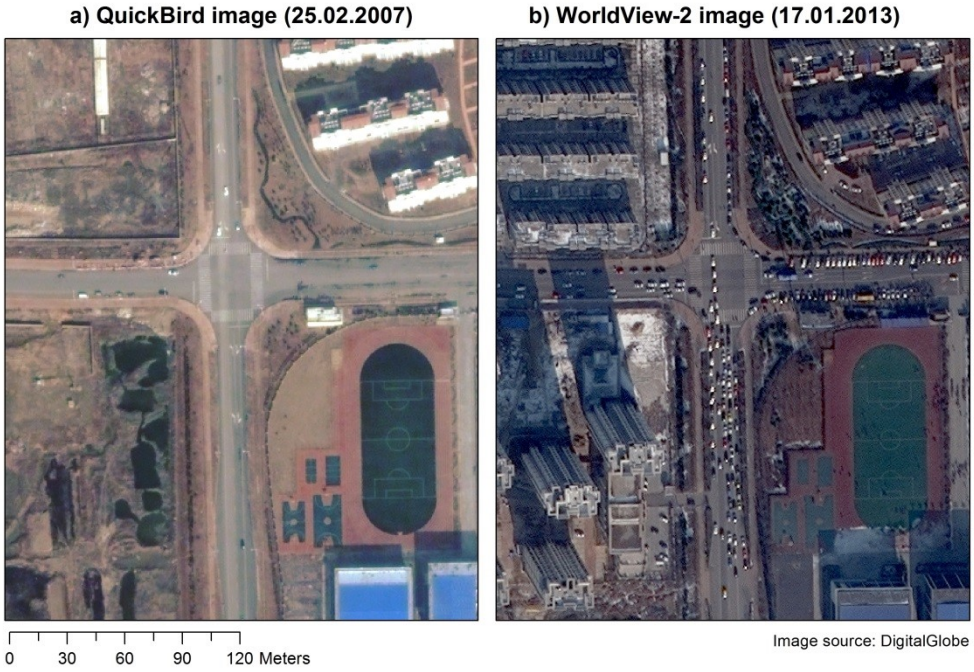

Fig. I-2: Examplary multi-temporal VHR image pair over Dongying, China. a) QuickBird image acquired on 25.02.2007, b) WorldView-2 image from 17.01.2013.

Fig. I-2 depicts an example of a multi-temporal image pair in a dynamic urban area of the city of Dongying, China for illustration of the framework for change detection (Fig. I-1). Obviously, a considerable amount of changes occurred on the ground (Ω_{Grd}) within the temporal interval of six years. Dependent on the application, several anthropogenic activities might be of interest, for example the construction of buildings and corresponding loss of other land cover or the changing traffic volume (i.e., the number of cars). However, different environmental conditions in terms of snow cover or different phenological states of the vegetation are also visible. Besides these differences affecting the Earth's surface, the atmospheric conditions (Ω_{Atm}) deviate in between the acquisition dates. Finally, differences in the acquisition system (Ω_{Sys}) induce changes in the image pair. The different sensors (i.e., QuickBird and WorldView-2) possess deviating sensor acquisition modes, i.e., different number of multispectral bands with slightly deviating wavelengths. Deviating sensor view angles cause

significantly different representation of vertical objects due to the parallax distortion, which particularly affects high-rise buildings. In addition, the intra-annual temporal difference of 39 days causes variations of the solar radiation incidence angle which manifests in different shadowing of vertical objects in the multi-temporal image pair.

In general, the taxonomy presented in Fig. I-1 can be customized to specific applications and available VHR images while irrelevant parts can be removed for proper modeling of the change detection problem and facilitation of a highly accurate change detection approach. However, especially for unsupervised and highly automated implementation of the change detection methodology, appropriate preprocessing methods must be applied in order to separate non-relevant radiometric differences and enable proper identification of changes that are relevant to the specific application (Bruzzone and Bovolo 2013).

2.2 Methods for change detection using VHR remote sensing

To date, numerous change detection techniques have been proposed based on optical remote sensing imagery with moderate and high spatial resolution (Coppin et al. 2004; Lu et al. 2004; Radke et al. 2005; Singh 1989). However, not all of these methods are capable of proper identification of changes due to the complexity of VHR remote sensing image analysis (Bruzzone and Bovolo 2013; Hussain et al. 2013; Lu et al. 2014a).

In general, the change detection workflow can be divided into four steps according to Fig. I-3. *First*, the multi-temporal remote sensing imagery must be preprocessed in order to mitigate differences in atmospheric conditions or to ensure exact spatial matching of the data. *Second*, the selected unit of analysis is commonly separated in pixel- and object-based approaches at the most general level. At this stage, the input features for change detection are derived and prepared based on the unit of analysis. *Third*, the comparison method for determination of changes can be applied depending on the availability of prior knowledge on the change detection task, which generally divides techniques into supervised and unsupervised methods. *Finally*, the resulting change map may either hold binary information on the occurrence of changes only or contain more detailed characteristics on the magnitude and/or type of changes. This step also includes the mandatory assessment of the change detection accuracy.

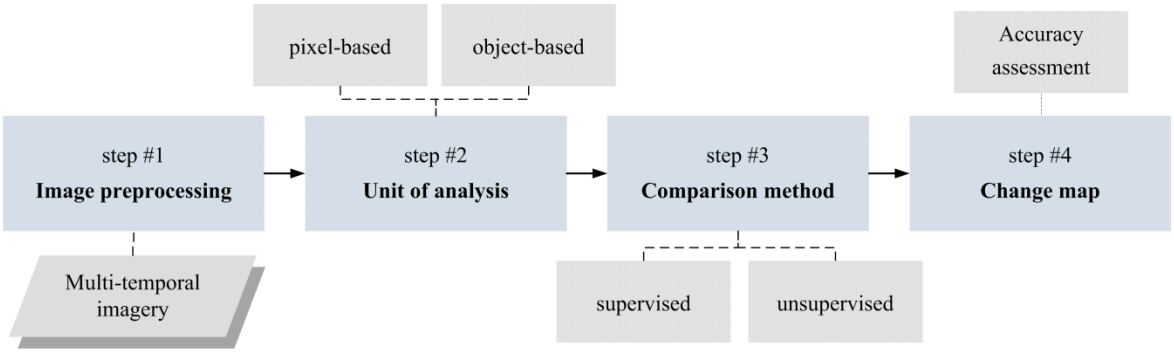

Fig. I-3 General change detection workflow (modified based on Tewkesbury et al. (2015)).

step #1 - Image preprocessing

According to Webb (2011), proper preprocessing of the input data is one of the most relevant tasks in any data analysis task. Preprocessing does not only increase the accuracy of change detection results, but also decreases the algorithmic complexity as well as computing time (Kantardzic 2011). Foremost, preprocessing applies to the multi-temporal remote sensing imagery which must be geometrically aligned accurately. In addition, differences in atmospheric conditions between the multi-temporal acquisitions due to changing weather conditions and variations in atmospheric transmittance can be mitigated by means of atmospheric correction. The importance of preprocessing with respect to change detection algorithms is highlighted by Radke et al. (2005), its particular relevance for unsupervised object-based methods is discussed by Bruzzone and Bovolo (2013).

step #2 - Unit of analysis

At the second stage of the change detection workflow (Fig. I-3), methods can be categorized according to the unit of analysis, which is most commonly separated in pixel- and object-based approaches (Hussain et al. 2013). Although object-based methods for change detection are recommended in general over pixel-based methods for VHR images in recent review studies (Hussain et al. 2013; Lu et al. 2014a), pixel-based approaches remain frequently proposed. Furthermore, many change detection techniques are almost identical and vary only by the unit of analysis (e.g., change vector analysis (CVA)) which results in a significant technical overlap between pixel- and object-based methods (Warner and Almutairi 2009) and any pixel-based change detection technique could be applied to objects and vice versa (Tewkesbury et al. 2015). However, in multi-temporal VHR data settings accurate change detection results based on pixel-based methods are mostly only possible with identical or very similar parameters of the acquisition system (Ω_{Sys}) and exact spatial matching of the data. By implication, object-based change detection methods offer the capability of mitigating differences in the

acquisition system (Ω_{Sys}) or inaccuracies of georeferencing by exploiting the object-based paradigm using appropriate object boundaries and object-based features (Blaschke 2010). Especially for VHR imagery, object-based methods are beneficial in general since the considered image objects largely consist of a set of pixels that are reasonably analyzed together (Blaschke 2010). Furthermore, object-based image analysis enables the utilization of additional features (e.g., first-order statistics, texture, or spatial context) for change detection besides pure reflectance values of individual pixels (Chen et al. 2012).

Dependent on the general principle of image object generation and consequent analysis, two major categories of object-based change detection approaches were consistently identified (with slightly different terminologies and few modifications) by the review studies of Chen et al. (2012), Hussain et al. (2013), Lu et al. (2014a), and Tewkesbury et al. (2015).

The *first category* describes image objects that are initially generated in each multi-temporal image separately and subsequently compared between acquisition dates. This technique enables the most accurate representation of objects in each individual image. In addition, also features with respect to the size or shape of objects can be compared for change detection. However, proper establishment of spatial and temporal links between image objects is challenging and may introduce so called sliver objects in case of polygon intersections (Tewkesbury et al. 2015). Early examples of separate image objects for change detection are Niemeyer et al. (2008) and Lefebvre et al. (2008), who employed QuickBird and aerial images for detection of land cover changes. More recently, Liu et al. (2018a), Xiao et al. (2017), and Zhang et al. (2017) performed change detection of building in VHR imagery based on separate segmentations while focusing on the spatio-temporal linkage of multi-temporal objects. A modified technique from this category of image object generation is the segmentation of one image and subsequent overlay on another image. For example, Listner and Niemeyer (2011) use this concept for change detection in multi-temporal QuickBird imagery.

The *second category* of change detection approaches is based on multi-temporal image object generation, where common object geometries are employed for all images. In general, these geometries are created by segmentation of a stack of multi-temporal images. This technique provides the advantage of image objects that are consistent in size and shape over time (Hussain et al. 2013). In addition, multi-temporal image objects are spatially consistent while highlighting key multi-temporal boundaries (Tewkesbury et al. 2015). However, differences due to acquisition conditions (Ω_{Acq}) (e.g., atmospheric conditions, sensor view angles etc.) between the images might blur the segmentation result (Chen et al. 2012). The first important research studies who analyzed changes in natural landscapes based on multi-temporal image objects were Bontemps et al. (2008), Desclée et al. (2006), and Stow et al. (2008). Doxani et al. (2012) employed multi-temporal objects for change

detection in urban areas, while Zhang et al. (2018b) extended this concept to a multi-temporal multi-scale representation of image objects for change detection. A special case of the second category is the utilization of vector polygons that are superimposed as multi-temporal image objects for change detection. In this context, it is crucial that the spatial scale of objects as well as the spatial reference matches the multi-temporal imagery (Tewkesbury et al. 2015). Frequently, this technique is used for map updating, where remote sensing imagery is used for the identification of changes based on existing vector polygon maps. Examples of such approaches are Bouziani et al. (2010) and Durieux et al. (2008), who detected building changes in VHR imagery due to construction and urban growth. Opposed to this, Sofina and Ehlers (2016) and Ye et al. (2017) employ this technique for building damage detection after an earthquake.

In addition, Chen et al. (2012) as well as Tewkesbury et al. (2015) include hybrid approaches that combine pixel- and object-based analysis as a special category for change detection. Examples from this special category are Lv et al. (2019) and Xiao et al. (2016), who employ object-based post-processing methods on initial pixel-based change detection results.

step #3 - Comparison method

According to the third step of the change detection workflow depicted in Fig. I-3, an important feature for the categorization of change detection algorithms is the availability of prior knowledge which separates supervised and unsupervised techniques. The choice of the change detection strategy in this regard is foremost dependent on the availability of reliable multi-temporal training samples in temporal accordance with the multi-temporal remote sensing imagery (Richards 2012). In addition, supervised methods classify the input data according to the provided labels while unsupervised techniques aim to identify the inherent structure of the data, which is a conceptually different approach in general (Webb 2011). Considering this fundamental methodological difference and in view of transferability and automatization, the necessity of collecting proper a priori information (i.e., labels for training) in case of supervised change detection techniques complicates this task. In contrast, the absence of a priori information in case of unsupervised change detection approaches requires proper preprocessing of the input data in order to supplant irrelevant radiometric differences for accurate detection of changes with respect to the specific application (see Section 2.1).

Review studies on change detection provide exhaustive lists and categorizations of supervised and unsupervised algorithms that can be employed for multi-temporal comparison of pixels or image objects. For example, Lu et al. (2004) distinguish 31 different techniques, and Bhagat (2012) discriminates 29 distinct methods for change detection. In order to provide a concise summary within the scope of this dissertation, the

most important and recently applied techniques are summarized into two major categories while less used methods are not described in detail.

The *first category* of comparison methods can be characterized as arithmetic and algebraic methods that compare the multi-temporal imagery based on radiometric differences between the images. These methods utilize spectral values of the image bands or derived indices, such as the Normalized Difference Vegetation Index (NDVI). The most straightforward way of change identification in this regard is image differencing, which reflects the magnitude of change by subtracting one image from another. Examples based on this techniques are Bovolo and Bruzzone (2007a), or Hall and Hay (2003). In particular for object-based change detection, the consideration of texture and spatial context extends the capabilities of arithmetic method for change detection (e.g., Klaric et al. (2013), Zhu et al. (2017)). For determination of a binary change map, these methods are frequently combined with unsupervised as well as supervised decision functions like thresholding (Falco et al. 2013), k-means clustering (Tang and Zhang 2017), the expectation–maximization (EM) algorithm (Li et al. 2017b), support vector machines (SVM) (Volpi et al. 2013), or deep learning methods (Cao et al. 2017; Mou et al. 2019). Another popular change detection method that can be attributed to algebraic methods is CVA. Besides its magnitude, this technique also enables the assessment of the type of change by the direction of the change vector (Bovolo and Bruzzone 2007b). CVA was frequently applied for both, pixel- and object-based change detection, as well as in terms of supervised and unsupervised comparison. For example, supervised CVA was employed in a pixel-based manner by Varshney et al. (2012), while Liu et al. (2018a) utilized this technique on objects. In addition, Thonfeld et al. (2016) implemented unsupervised pixel-based CVA, and Chen and Chen (2016) proposed this method in an object-based way. A recent example using a combination of different methods is Bullock et al. (2019), who employ an ensemble of algorithms for change detection.

The *second category* of comparison techniques for change detection is based on classification methods, which either classify the multi-temporal image stack directly or perform a comparison of multi-temporal classification maps in terms of post-classification change detection (Hussain et al. 2013). For this purpose, a variety of classification approaches was utilized in literature. For example, Xian et al. (2009) and Brunner et al. (2010) introduced threshold-based classification methods on pixels as well as image objects for multi-date direct classification. Other commonly used classification techniques to mention in this regard are SVM (e.g., Li et al. (2010), Shah-Hosseini et al. (2017)) or random forest (RF) (e.g., Li et al. (2015)). Recently, also deep learning was applied for change detection based on multi-date direct classification (Zhang et al. 2019b; Zhang et al. 2016). These techniques were equally employed for post-classification change detection (e.g., Lefebvre and

Corpetti (2017), Ye et al. (2016)). In addition, other approaches like rule-based classification (Zhou et al. 2008), or clustering (Li et al. 2017a) were proposed for classification based change detection.

step #4 - Change map

The fourth and final step of the change detection workflow depicted in Fig. I-3 is the change map, which directly results from the previous step. Dependent on the change detection approach, the change map may hold binary information on the occurrence of changes or more detailed characteristics on the magnitude and type of changes. The confusion or error matrix, which opposes classified against reference areas (i.e., pixels or objects) for each class, is most frequently established for mandatory accuracy assessment (Congalton 1991). Common measures of accuracy which can be delineated from the confusion matrix are, among others, the overall accuracy (Congalton 1991), Cohen's kappa coefficient κ (Foody 2002), the F-score (Van Rijsbergen 1979), the True-Skill-Statistics (TSS) (Allouche et al. 2006), or Receiver Operator Characteristics (ROC) as well as Precision-Recall (PR) curves (Davis and Goadrich 2006).

2.3 Fields of application for change detection based on VHR data in cities

The increasing availability of VHR remote sensing imagery (Section 2) in combination with an appropriate methodology for the reliable identification of changes (Section 2.2) enables a wide range of fields of application for change detection in cities. First, as proposed by Bruzzone and Bovolo (2013) in Fig. I-1, changes that occurred on the ground (Ω_{Grd}) can be analyzed for urban applications. Second, applications that combine changes of land cover from the first category with ancillary data (e.g., statistical data, census counts, etc.) allow a more detailed characterization of intra-urban dynamics and their geographical implications. Finally, a high potential for the development of new fields of multi-temporal applications exists for both categories and will be outlined based on existing mono-temporal studies in this regard.

First, applications based on changes of land cover in urban areas have been proposed in literature since the early 1970s (Dueker and Horton 1972). Based on remote sensing imagery with high spatial resolution, several case studies were developed at different spatial scales from city level to nationwide analyses of urban growth (e.g., Huang et al. (2017), Li et al. (2015), Taubenböck et al. (2012), Taubenböck and Wiesner (2015), Wang et al. (2012a)). In addition, first global multi-temporal applications on urban growth were published based on HR data (Esch et al. 2018; Liu et al. 2018b). With regard to VHR remote sensing imagery, similar applications on urban growth were established (Xiao et al. 2016). However, due to a lack of VHR data at medium and small

spatial scales, these studies were only conducted locally (mostly for single VHR scenes) but at high level of detail (i.e., at the level of individual buildings) (Wang et al. 2018b; Xiao et al. 2017). To value the unique capabilities of VHR data in terms of classification of the intra-urban morphologic configuration of urban areas, several more detailed urban growth studies have been developed. For example, Gruebner et al. (2014) and Kit and Lüdeke (2013) conducted the multi-temporal detection and assessment of informal settlements as a special morphologic appearance in urban areas, while Lefebvre and Corpetti (2017) aim at the identification of morphologic transformations (i.e., intra-urban structural changes). In contrast to urban growth, Thompson and de Beurs (2018) tracked the removal of buildings and consequent shrinking of cities. Another application in this regard is monitoring of impervious surfaces, which was demonstrated by Zhang and Huang (2018). Concerning vegetation cover, Handayani et al. (2018) and Wang et al. (2018a) explored the multi-temporal dynamics of urban green based on VHR remote sensing data. Another important and well-established field of application with regard to changes of land cover based on VHR remote sensing is the detection of damages as a result of a natural disaster. Frequently, multi-temporal VHR imagery was utilized for the detection of damages to the built environment after an earthquake (e.g., Brunner et al. (2010), Falco et al. (2013), Li et al. (2010), Yeom et al. (2017)) or a tsunami (e.g., Olsen et al. (2013), Pesaresi et al. (2007), Shah-Hosseini et al. (2017), Tanathong et al. (2008)). Other applications of damage detection by means of VHR imagery were conducted in the aftermath of landslides (Gong et al. 2008; Park and Chi 2008), typhoons (Gueguen and Hamid 2016), or different situations of armed conflicts (Klonus et al. 2012; Knoth and Pebesma 2017; Kranz et al. 2017).

Additional potential for applications based on changes of land cover lies in the analysis of the intra-urban morphologic configuration, where the mono-temporal classification of urban functional zones based on VHR remote sensing imagery was conducted at different levels of detail (Li et al. 2016; Lin et al. 2016; Zhang et al. 2018a). Similar classification approaches targeting urban morphologic structure and function were implemented by Li et al. (2018) as well as Luo et al. (2019) in the application context of urban ecology. Huang et al. (2013) estimated the mono-temporal distribution as well as volume of urban green, whereas a detailed typology of the urban vegetation was performed by e.g., Rougier et al. (2016) and Tigges et al. (2013). In this context, the carbon storage (Schreyer et al. 2014) as well as the water consumption (Di et al. 2019) of urban vegetation were investigated based on VHR data. The urban morphology was also utilized for wind speed estimation in the context of urban meteorology (Kent et al. 2019), which is also a relevant parameter for noise modeling (Nega et al. 2012) as well as in urban temperature studies (Bechtel et al. 2015; Stewart and Oke 2012). In this context, the urban heat island effect was investigated by several studies employing information from mono-temporal VHR data (Alavipanah et al. 2018; Yu et al. 2018; Zhang and Sun 2019). The assessment of the potential for heat

supply (Geiß et al. 2011) or photovoltaic energy generation (Santos et al. 2014) was also proposed based on VHR data. Finally, the potential of VHR imagery for vehicle detection and traffic monitoring was demonstrated by Kopsiaftis and Karantzalos (2015) as well as Li et al. (2019).

Second, the category of multi-temporal applications based on VHR data that combine changes of land cover with ancillary data have hardly been investigated recently (Si Salah et al. 2019). In this context, Veljanovski et al. (2012) explored multi-temporal population estimation based on multi-temporal VHR imagery and census data in informal settlements.

With regard to applications that combine VHR land cover from remote sensing with ancillary data, several mono-temporal studies exist. For example, the dwellings of a refugee camp were determined by Aravena Pelizari et al. (2018) and Spröhnle et al. (2017), whereas the water consumption of informal settlements was estimated by Niebergall et al. (2008). The analysis of vulnerability and risk against natural disasters is an important field of application based on VHR remote sensing and was demonstrated in the context of earthquakes (Geiß et al. 2015a; Pittore and Wieland 2013), tsunamis (Eckert et al. 2012; Römer et al. 2012), or flood events (Sowmya et al. 2015; Taubenböck et al. 2011). Several socio-economic parameters have been related to the urban morphology based on mono-temporal VHR satellite images, for example the economic status of residents (Wurm and Taubenböck 2018), the value of housing (Taubenböck et al. 2009), land rent (Wu et al. 2019), or incidents of crime (Patino et al. 2014; Sathyakumar et al. 2018). Quality of life in general was estimated using VHR remote sensing data by Cabrera-Barona et al. (2016), among others. Finally, population estimation is an important field of application in urban areas where VHR remote sensing was frequently utilized (Biljecki et al. 2016; Taubenböck et al. 2007; Wu and Murray 2005).

Although this list of studies with regard to the two categories of applications may not be exhaustive, it becomes evident that the capabilities of multi-temporal VHR imagery have not been exploited in all possible fields of application. The wealth of mono-temporal studies demonstrates the potential for multi-temporal analysis using VHR remote sensing data and appears feasible in many regards (Lu et al. 2004; Miller and Small 2003; Qin et al. 2016). Against this background, a multi-temporal analysis could introduce additional relevance and a different perspective to the investigated topic of urban geography and enable new fields of application.

3. Study area and background

The global trend of urbanization is particular dynamic in China, which possesses among the highest urbanization rates worldwide (Seto et al. 2011; Taubenböck et al. 2014). Based on 1.4 billion inhabitants in 2017, China is the most populous country of the world (United Nations 2017). However, area-wide, homogeneous, and up-to-date data and derived information on urbanization is still lacking in many regards. In addition, the political system of China and urban planning provide a globally unique framework condition for urban development (Gar-on Yeh and Wu 1999; Wu 2015). In this regard, several studies have investigated urban growth in China based on remote sensing data with medium and high spatial resolution at different spatial scales, e.g., individual cities (Chen et al. 2014; Li et al. 2015), urban agglomerations (Tan et al. 2005; Taubenböck et al. 2014) as well as at national scale (Wang et al. 2012a; Xiao et al. 2014). However, research studies based on remote sensing beyond the application of urban growth are still in the minority. In accordance with the expansion of urban areas, population numbers have generally increased massively, especially in relation to the urban population (Ruibo and Linna 2013). As shown in Fig. I-4, for the first time around the year 2010, more people were living in urban areas than in rural areas in China.

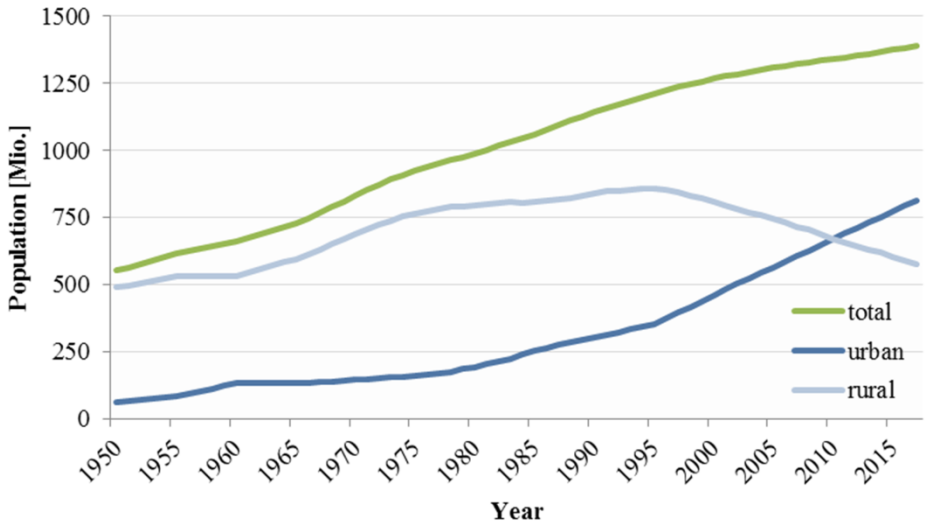

Fig. I-4 Urban and rural population development in China from 1950 to 2017 according to National Bureau of Statistics of China (2018).

In general, the majority of the studies related to urban remote sensing are focused on large Chinese cities like Beijing, Tianjin, Shanghai, Shenzhen, or Guangzhou, as well as the well-known urban agglomerations of the Pearl River Delta, the Yangtze River Delta, or the Beijing-Tianjin-Hebei region (Zhang et al. 2018c). Besides its five megacities with more than 10 million inhabitants (i.e., Shanghai, Beijing, Tianjin, Guangzhou, and Shenzhen), China has more than 240 cities with population of more than 0.5 million inhabitants according to the National Bureau of Statistics of China (2018). Especially the smaller cities of the eastern provinces as well as along the coasts of China possess sustained rapid development and very high dynamics (Wen et al. 2016), particularly in the vicinity of dynamic urban agglomerations such as the Pearl River Delta, the Yangtze River Delta, the Beijing-Tianjin-Hebei region, the middle reaches of the Yangtze River around Wuhan, and the Chengdu–Chongqing region (Shi et al. 2019).

A region of special importance and interest with respect to urbanization dynamics is Shandong province and in particular the Yellow River Delta (YRD). The YRD is situated in the North of Shandong province, located around 300 km southeast of the Beijing-Tianjin-Hebei region at coast of the Bohai Sea (Fig. I-5). The YRD is part of the second largest oil field of China, the Shengli oil field, which developed massively from the construction of the first oil wells in the 1970s (Kuenzer et al. 2014). In the course of this development, the city of Dongying was established in 1983 as a harbor for oil related industries and their workers (Ottinger et al. 2013). Recently, the region also became an important industrial zone for salt production and aquaculture (Ottinger et al. 2016; Wohlfart et al. 2016a). The YRD together with its main city of Dongying is considered as one of the key regions in China's future economy (Jiang et al. 2011) and was promoted to a special economic area at the national level called the Yellow River Delta High-efficient Eco-Economic Zone (Wang et al. 2016a). Thus, a fast economic development has taken place during the past decades (Wohlfart et al. 2016b) and urbanization and industrialization in the YRD make it one of the fastest growing delta regions worldwide (Ottinger et al. 2013), which introduces several ecological challenges and threatens the natural resources and ecology of the area. These rapid developments led to environmental degradation in general (Wohlfart et al. 2016a), especially with respect to the valuable wetland ecosystems which suffer from fragmentation (Liu et al. 2014), loss, and degradation (Wang et al. 2012b). In this regard, two nature reserves have been established by the government in the beginning of the 1990s aiming at the restoration and conservation of the unique wetland ecosystem (Cui et al. 2009), but which have been proven to be under severe violation by construction activities, especially in relation to oil extraction (Kuenzer et al. 2014). In addition, the exploitation of oil puts heavy pressure on local natural ecosystems and causes incalculable damage to the environment (Wang et al. 2016a). Increasing domestic and industrial water demand leads to water shortage (Xu et al. 2002) and water pollution (Baosheng et al. 2004) of

the Yellow River. Finally, as revealed by Zhang et al. (2019a), industry-related extraction of oil and groundwater as well as the construction of buildings, railways, and roads cause severe land subsidence in the whole region of the YRD.

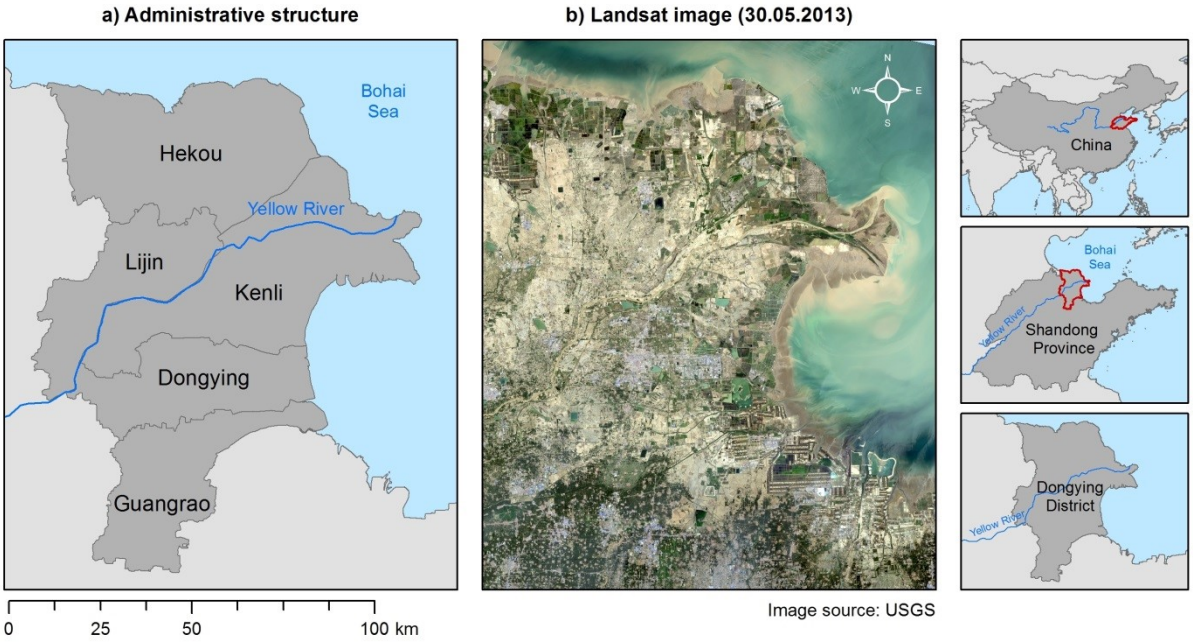

Fig. I-5 Study area of the Yellow River Delta: a) Administrative structure within Dongying district, b) Landsat image of the Yellow River Delta.

4. Objectives and aims of this dissertation

In view of rapidly changing urban environments and the increasing availability of multi-temporal remotely sensed VHR data, this thesis aims at the development of a change detection approach in an urban environment using data from different sensors and consequent exemplary application of the assessment of the ghost city phenomenon in the context of urban geography. In this regard, three specific goals are defined within the scope of this dissertation.

1) Development of a change detection approach for VHR data from different sensors

Based on the envisaged utilization of optical VHR remote sensing data for identification of land cover changes in urban areas, an object-based approach is developed within this thesis for change detection. The approach focuses on the detection of newly constructed individual buildings, which represent one of the most distinctive changes of land cover regarding urban morphology in the context of urban growth. In addition, suitable capabilities for handling data with differences in the acquisition system, possibly from different sensors, are included and examined in order to enable flexibility with regard to the input data and thus, full utilization of the growing amount of VHR imagery available.

Technically, the developed approach is based on multi-temporal image objects, which are classified into changed and unchanged buildings by means of clustering after exhaustive preprocessing of the imagery and associated object-based difference features. Due to its unsupervised implementation, the approach offers high potential for automation in order to enable efficient processing of the increasing amount of VHR imagery. In this context, different unsupervised clustering techniques are evaluated. The change detection approach is implemented and evaluated in an experimental test site in the city of Dongying, China.

2) Evaluation of transferability with regard to the distribution of classes

The most important characteristic of a versatile and flexible change detection technique is its transferability to a different geographical area based on a deviating data setting. For this reason, the transferability of the developed approach is evaluated with particular focus on the nature and effects of class distribution and class imbalance. The distribution of classes is unknown a priori and varies with the type of application, geographical setting, temporal scale, as well as the spatial extent of the remote sensing imagery of the binary change detection setting in urban environments.

For this purpose, a diagnostic framework for evaluation of model transferability in any two-class classification (or change detection) problem is developed. It is based on the methodological approach from the first objective in this thesis and examines different partitioning clustering techniques that were found suitable for unsupervised change detection. This framework is applied to two study sites with different characteristics of the built-up structure as well as divergent temporal evolution: the city of Dongying, China and the city of Munich, Germany.

3) Exemplary application of the assessment of the ghost city phenomenon

Within the scope of this dissertation, the examination of a background aspect of urban growth in China is chosen as an exemplary application in the context of urban geography. In contrast to the well-known trends of rapid urbanization, population growth, and migration towards urban areas in China, a recently emerging and rarely studied side effect is the ghost city phenomenon. Although several lists and rankings identify ghost cities across China, their assessment is largely inconsistent and more objective and scientific profound methodological frameworks are required. A ghost city is commonly defined as a new urban development running at severe undercapacity, a place which houses drastically fewer people and businesses than there is available space for (Shepard 2015). As opposed to ghost estates in other countries around the globe which mostly arise from abandoned old residential developments, a unique feature of Chinese ghost cities is that completed new buildings are affected. However, a ghost city clearly contradicts the envisaged goal of sustainable urbanization since it induces a significant consumption of land, resources and energy while it contributes to increasing air emissions and creates unnecessary construction and demolition waste (He et al. 2016).

For investigation of the ghost city phenomenon, a conceptual framework for the assessment of its presence or absence is designed based on VHR imagery and census counts, which relates changes of the urban morphology to the temporal evolution of population. In this manner, undercapacity with respect to the residential population as one of the key characteristics of a ghost city is replicated. Multi-temporal VHR images represent a suitable basis for establishment of a 4d functional city model, which is subsequently employed for estimation of available living space of residential buildings (i.e., population capacity). Population capacity estimates are opposed to actual permanent residential population (which is available in terms of ancillary data from census) for the assessment of the ghost city phenomenon in the city of Dongying, China.

5. Thesis structure

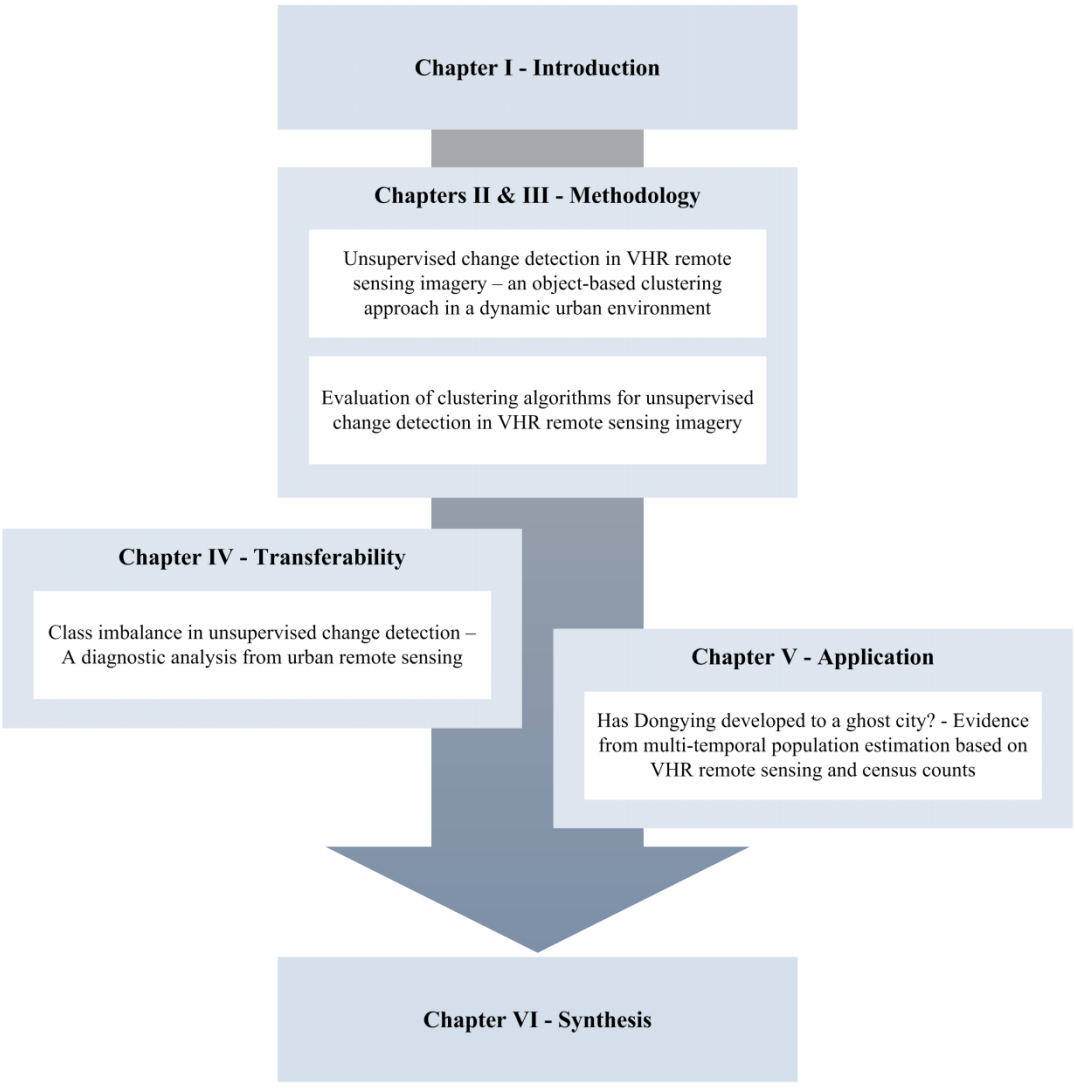

Fig. I-6 Visualization of the structure of this thesis.

This cumulative thesis is structured as displayed in Fig. I-6. Chapter I provides a general introduction to the topic of change detection using remote sensing imagery with particular focus on optical VHR data and urban environments. In addition, the motivation for selecting the study area as well as the objectives and aims of this dissertation are outlined. In Chapters II - V, the core research of the cumulative thesis is presented in terms of four stand-alone manuscripts that have been published in or have been submitted to international, peer-reviewed journals.

Chapter II *Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment.* Published in the International Journal of Applied Earth Observation and Geoinformation 54 (2017), 15–27.

Chapter III *Evaluation of clustering algorithms for unsupervised change detection in VHR remote sensing imagery.* Published in the Proceedings of the 2017 Joint Urban Remote Sensing Event (JURSE), IEEE Xplore (2017), 1-4.

Chapter IV *Class imbalance in unsupervised change detection – A diagnostic analysis from urban remote sensing.* Published in the International Journal of Applied Earth Observation and Geoinformation 60 (2017), 83–98.

Chapter V *Has Dongying developed to a ghost city? - Evidence from multi-temporal population estimation based on VHR remote sensing and census counts.* Computers, Environment and Urban Systems 78 (2019), 101372.

The methodology for change detection in this thesis is introduced in Chapter II, while the peer-reviewed conference proceeding in Chapter III presents a methodological extension. These two chapters contribute to the first research objective of this thesis. Chapter IV is allocated to the second objective of this work and provides an analysis on the distribution of classes in this context, which is a crucial factor with respect to transferability of the proposed approach. With regard to the third objective of this dissertation, Chapter V introduces the assessment of the ghost city phenomenon as an exemplary application from urban geography using multi-temporal VHR remote sensing imagery and the developed change detection methodology as a central element of the proposed framework for the ghost city assessment in the Chinese city of Dongying. Chapter VI concludes this thesis and provides a synthesis and discussion on the major outcomes. Finally, an outlook on the specific research topics of this work as well as from a more general perspective regarding monitoring of urban areas using VHR remote sensing imagery is provided.

Chapter II. Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment

International Journal of Applied Earth Observation and Geoinformation 54 (2017), 15–27.

Tobias Leichtle[a, b, c], Christian Geiß[b], Michael Wurm[b], Tobia Lakes[c], and Hannes Taubenböck[b]

[a] Company for Remote Sensing and Environmental Research (SLU), Kohlsteiner Straße 5, 81243 München, Germany

[b] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchner Straße 20, 82234 Weßling, Germany

[c] Humboldt-Universität zu Berlin, Geography Department, Rudower Chaussee 16, 12489 Berlin, Germany

http://dx.doi.org/10.1016/j.jag.2016.08.010
Received 17th May 2016; revised 08th August 2016; accepted 31st August 2016.

Abstract

Monitoring of changes is one of the most important inherent capabilities of remote sensing. The steadily increasing amount of available very-high resolution (VHR) remote sensing imagery requires highly automatic methods and thus, largely unsupervised concepts for change detection. In addition, new procedures that address this challenge should be capable of handling remote sensing data acquired by different sensors. Thereby, especially in rapidly changing complex urban environments, the high level of detail present in VHR data indicates the deployment of object-based concepts for change detection. This paper presents a novel object-based approach for unsupervised change detection with focus on individual buildings. First, a principal component analysis together with a unique procedure for determination of the number of relevant principal components is performed as a predecessor for change detection. Second, k-means clustering is applied for discrimination of changed and unchanged buildings. In this manner, several groups of object-based difference features that can be derived from multi-temporal VHR data are evaluated regarding their discriminative properties for change detection. In addition, the influence of deviating viewing geometries when using VHR data acquired by different sensors is quantified. Overall, the proposed workflow returned viable results in the order of κ statistics of 0.8 to 0.9 and beyond for different groups of features, which demonstrates its suitability for unsupervised change detection in dynamic urban environments. With respect to imagery from different sensors, deviating viewing geometries were found to deteriorate the change detection result only slightly in the order of up to 0.04 according to κ statistics, which underlines the robustness of the proposed approach.

1. Introduction

The broader availability of very-high resolution (VHR) remote sensing imagery, especially in the temporal domain, requires new concepts for change detection (Bruzzone and Bovolo 2013). Besides very-high resolution aerial and SAR imagery, space-borne optical VHR images with spatial resolution of 1m or less are acquired by satellite systems such as IKONOS, QuickBird, GeoEye, WorldView or Pléiades. After around 15 years of VHR data collection by these systems, the analysis of huge amounts of data necessitates highly automated unsupervised methods for change detection (Bruzzone and Bovolo 2013; Hussain et al. 2013). Regarding rather traditional medium (e.g., MODIS, AVHRR) and high resolution (e.g., Landsat, SPOT) sensors, a broad range of

efficient, mostly pixel-based methods for change detection exist (Coppin et al. 2004; Lu et al. 2004; Singh 1989). Due to the high level of detail present in VHR imagery, change detection becomes more complex and traditional (pixel-based) methods are considered less effective (Bruzzone and Bovolo 2013; Hussain et al. 2013). This becomes even more crucial in urban environments where heterogeneous surface materials are concentrated in a highly dynamic manner and thus, VHR remote sensing images entail a wealth of details. In the context of such highly dynamic urban environments, object-based approaches for change detection are considered more suitable (Hussain et al. 2013; Lu et al. 2014a). Especially in case of multimodal data sources, e.g., VHR imagery acquired by different sensors with deviating viewing geometries, position of the sun, among many others, exact geometric, spectral and radiometric matching of the multi-temporal images is almost impossible and object-based methods appear more appropriate and effective (Tewkesbury et al. 2015).

For object-based change detection of multi-temporal VHR data, basically three different methodological approaches are reported in literature (Chen et al. 2012; Hussain et al. 2013; Lu et al. 2014a): (1) direct comparison of individual segmentation objects, (2) comparison of classified objects, and (3) simultaneous segmentation and classification of multi-temporal objects. Methods from the *first category* usually perform segmentation of multi-temporal imagery separately and subsequently compare the resulting image objects (Chen et al. 2012). This comparison can be based on either geometrical properties, spectral information, and/or other attributes of the individual image objects (Hussain et al. 2013). The major drawback of methods from this category is that inconsistent geometries of the resulting objects must be detected and handled (Blaschke 2005). Approaches from the *second group* classify image objects in the individual multi-temporal images in order to derive thematic "from-to" changes (Chen et al. 2012). This technique refers to post-classification comparison in pixel-based change detection (Lu et al. 2004; Singh 1989) and is probably the most commonly used methodology for object-based change detection (Hussain et al. 2013). Results of these methods are intuitive and straightforward due to the thematic information content (i.e., "from-to" changes) but the methodology and accuracy of the initial classifications is crucial due to error propagation (Chen et al. 2012). Furthermore, geometric inconsistencies can deteriorate the result in case of deviating object geometries (Blaschke 2005). As a solution to the limitations of the above-mentioned approaches, techniques from the *third category* delineate multi-temporal objects from all scenes simultaneously and generate common image objects for all acquisitions (Hussain et al. 2013). Alternatively, mutual image objects from an external data source (e.g., from Volunteered Geographic Information or existing vector maps) can be utilized. Especially for object-based change detection approaches employing multimodal data sources as presented in this study, the concept of multi-temporal objects is highly beneficial due to consistency in size, shape and location coordinate of image objects over time (Chen et

al. 2012; Hussain et al. 2013). Furthermore, Tewkesbury et al. (2015) conclude that multi-temporal objects are likely the most robust analysis unit for object-based change detection.

The technique of multi-temporal objects has been applied successfully for change detection of VHR remote sensing imagery delineating changes of land cover in urban environments (e.g., Doxani et al. (2012), Im et al. (2008), Yang et al. (2015)). A special application of change detection in urban areas is damage detection of buildings (i.e., changes that have occurred due to natural hazards such as earthquakes or tsunamis), which is conducted by Al-Khudhairy et al. (2005) using pan-sharpened IKONOS data, while Chen and Hutchinson (2007) deploy correlation analysis on panchromatic QuickBird imagery. Another specific application of change detection with respect to buildings is map updating where (usually) mono-temporal VHR data are utilized for change detection based on outdated (vector) map data. For example, Bouziani et al. (2010) present a supervised change detection approach where existing building geometries are used for learning and classification of IKONOS and QuickBird images, whereas Matikainen et al. (2010) employ VHR aerial imagery and map data for building detection and rule-based change detection. Furthermore, several studies utilizing multi-temporal remote sensing images rely on unsupervised methods for change detection of buildings (Huang et al. 2014; Klonus et al. 2012; Tang et al. 2013). All above-mentioned studies utilize VHR remote sensing imagery from a single sensor with similar viewing geometries. A recent study that presents a change detection approach for imagery from different sensors is Wang et al. (2015), where data from QuickBird and WorldView are simultaneously cross-sharpened in order to detect changes robustly using an unsupervised workflow. However, this particular study presents change detection without focus on individual buildings.

In contrast to the above-mentioned studies, this paper uniquely combines the concept of multi-temporal objects with an adequate representation of object-based features and proposes a novel unsupervised object-based change detection approach using VHR imagery in a dynamic urban environment at the building-level. Furthermore, a very heterogeneous data base is considered, incorporating remote sensing imagery from different sensors with significantly deviating acquisition parameters. In this study, monitoring of urban growth is demonstrated by the example of the Chinese city of Dongying as experimental site. The objectives of this work are as follows: i) development of an unsupervised object-based change detection approach using multi-modal VHR imagery for monitoring of urban growth, ii) quantification of the influence of deviating viewing geometries of optical VHR satellite systems on different platforms and iii) evaluation of several object-based feature sets of different characteristics in order to indicate the basic suitability of distinct types of features for identification of changed buildings.

2. Materials

2.1 Study area

The change detection approach is applied to the Chinese city of Dongying, located in the Yellow River Delta (YRD). The YRD harbors the Shengli Oil Field, which is the second-largest oil deposit in China (Kuenzer et al. 2014). After construction of the first oil wells in the 1960s, the city of Dongying was established in 1983 in order to meet the needs for industrial and residential areas of the oil industry. The YRD together with the city of Dongying is one of China's key regions in terms of economic development (Wohlfart et al. 2016b) and thus, the city itself as well as the whole YRD have undergone a rapid economic development during the past four decades (Kuenzer et al. 2014). This is accompanied by highly dynamic urbanization and industrialization, which is well visible in the main city of the YRD, Dongying, where the population already exceeded 800,000 inhabitants within its relatively short history of only 30 years (Editorial Committee of Dongying Statistical Yearbook 2013). According to Ottinger et al. (2013), the YRD is one of the fastest growing deltas worldwide, which does not only affect urban development of the city of Dongying, but also possibly threatens surrounding natural resources.

The experimental site is a relatively flat area without considerable topography and comprises about 550 buildings for experimental evaluation of the object-based change detection analysis. The distribution of classes (i.e., changed against unchanged buildings) is quite balanced, which is characteristic for the complete city of Dongying as well as for other dynamic Chinese cities with respect to the 6-year period covered by the VHR optical imagery (Fig. II-1). Various types of building changes are present within the experimental site, e.g., from bare soil to buildings, from water bodies to buildings, building conversion, etc., whereas the majority of changes correspond to newly built buildings (ca. 80% change from bare soil, ca. 15% change from water bodies) and only a minority (ca. 5%) of changed buildings accounts for building conversion (building conversion is only present in terms of building reconstruction, i.e., demolition and new construction, within the experimental site). For this reason, changed buildings are defined as newly built buildings in this study. Furthermore, different kinds of building occupancies are included, such as industrial usage, shopping malls or residential buildings of different sizes and shapes. The buildings within the experimental site exhibit a variation in building footprint area of 80 to 18800m² and a height range of few meters to 60m (mean height per building based on building footprint and nDSM).

2.2 Data sets

Launched in 2009, WorldView-2 is the first commercial satellite providing VHR imagery with up to 8 multispectral bands. The pan-sharpened image products possess slightly better spatial resolution of 50cm compared to QuickBird. The employed 4-band multispectral (blue, green, red, NIR – 16bit each) WorldView image was acquired on January 17, 2013 with 22.5° mean off-nadir view angle. This image is referred to as t_0.

QuickBird offered the highest spatial resolution in Earth observation (EO) technology when it was launched in 2001 and was one of the most successful VHR imaging missions until the end of the mission in late 2014. The sensor provides pan-sharpened image products at 60cm spatial resolution comprising multispectral bands of blue, green, red, and near infrared (NIR) with radiometric resolution of 16bit each. The image used in this study was acquired on February 25, 2007 with 14.2° mean off-nadir view angle. This image will be further treated as t_0-1.

Both optical data sets are available as standard image products, i.e., radiometrically corrected, sensor corrected, and projected to the Earth's surface.

In addition, height information is available by means of a DSM, which is processed by automatic stereo matching of three panchromatic images acquired by the Pléiades satellites. The raw Pléiades panchromatic imagery was acquired with a spatial resolution of 50cm in a tri-stereo configuration on October 18, 2013. Thus, the DSM corresponds to t_0 (WorldView imagery), since there are no inconsistencies during the nine month offset period within the experimental site. The resulting height product (Elevation4 DSM) possesses a pixel spacing of 4m. This data set is georeferenced by the data provider and spatially matches the VHR imagery. Therefore, no further geometric correction was needed.

The optical remote sensing imagery as well as an excerpt of the DSM covering the experimental site are depicted in Fig. II-1.

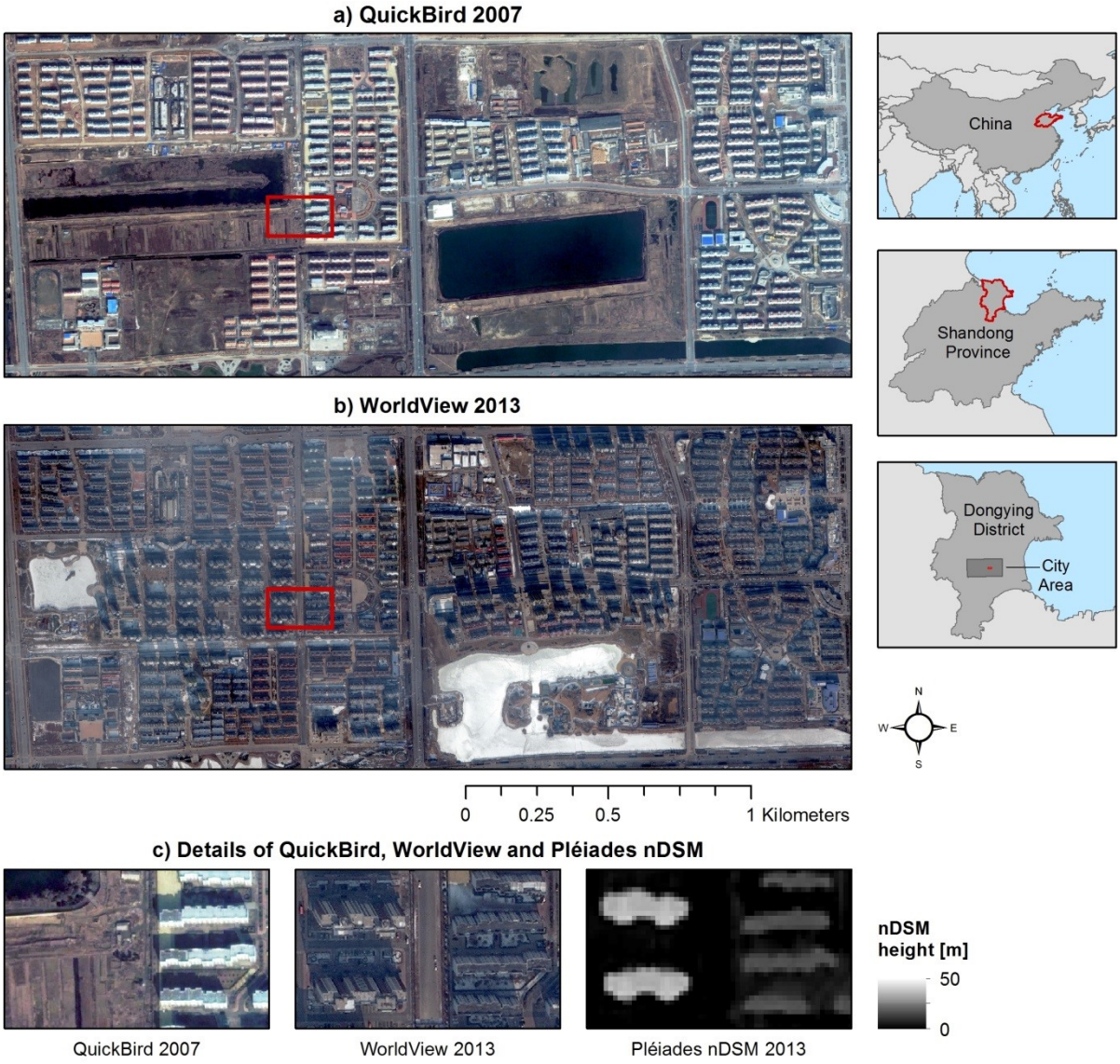

Fig. II-1 Study area and available data: a) QuickBird data 2007 (t_0-1), b) WorldView data 2013 (t_0), c) Detailed views of QuickBird, WorldView and Pléiades nDSM data, respectively.

3. Methodology

The proposed workflow for object-based unsupervised change detection is depicted in Fig. II-2. The initial preprocessing steps of optical imagery are introduced in Section 3.1.1, while the generation of object geometries based on a DSM is described in Section 3.1.2. Calculation and preparation of object-based difference features are explained in detail in Section 3.2. In Section 3.3, principal component analysis including a tripartite procedure for selection of relevant components is conducted. The change detection process in terms of clustering is presented in the subsequent section (3.4), where the widely used k-means algorithm (Hartigan and Wong 1979;

Jain 2010) is employed. Compulsory evaluation and comparison of the results is described in the final section of this chapter (3.5).

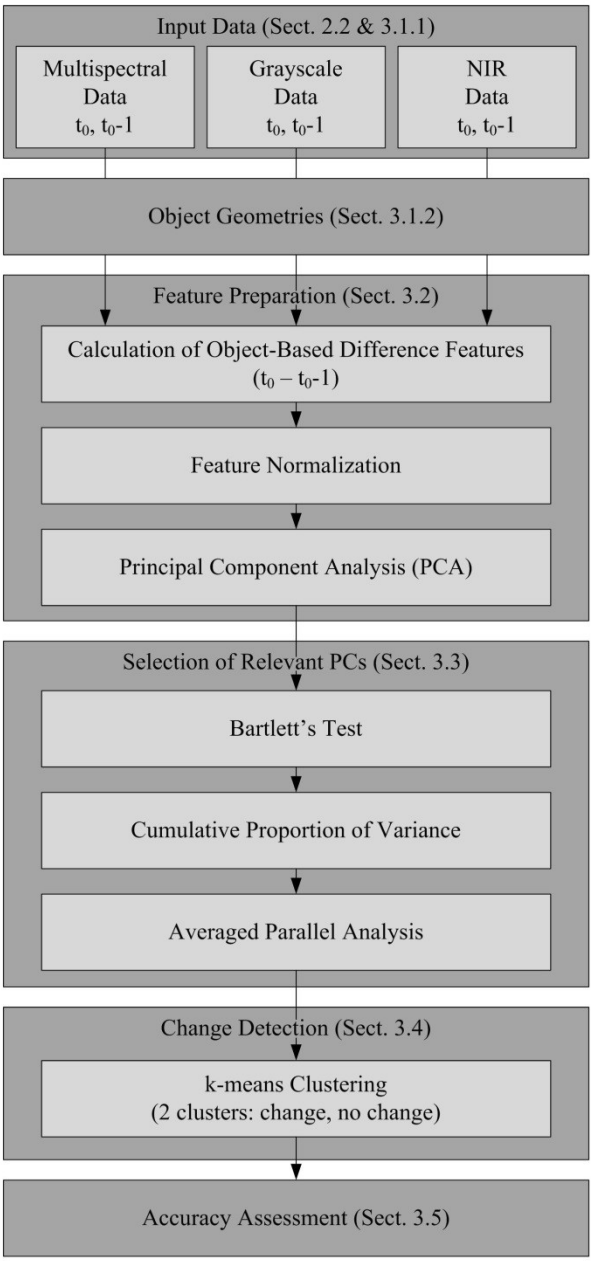

Fig. II-2 Overview of the proposed workflow.

3.1 Preprocessing

3.1.1 Preprocessing of optical imagery

The preprocessing procedure is of particular importance since insufficient radiometric correction and adjustment as well as poor geometric co-registration may lead to differences between the images that indicate change which are not associated to changes on the ground. Although the optical remote sensing images of QuickBird and WorldView are delivered as radiometrically corrected products by the data provider, further correction is needed due to influence of haze and smog. Other issues are deviating sensor viewing geometries as well as low solar illumination angles, which are especially present in the WorldView (t_0) imagery and result in severe tilting of buildings and long shadows (Fig. II-1). Further radiometric correction was accomplished by the ATCOR (Atmospheric and Topographic Correction) toolbox (Richter 1996). In order to finally adjust the images radiometrically, band-by-band histogram matching was applied, i.e., each pair of the multispectral bands (blue, green, red, NIR) was adjusted separately. The VHR images were properly co-registered geometrically using 60 ground control points (GCPs) based on the street network, since objects above ground (e.g., buildings) are influenced by deviating sensor viewing geometries. During geometric adjustment, the spatial resolution of the QuickBird (t_0-1) data was resampled to 50cm in order to match the geometric properties of WorldView (t_0) data. The resampling utilizes the nearest neighbor algorithm and entails slight oversampling due to reduction of spatial resolution. Mean RMSE (root mean square error) of co-registration was 1.5 pixels (i.e., 0.75m) and a common coordinate system (UTM Zone 50N) was selected.

3.1.2 Generation of object geometries

The extraction of building footprints is considered as an interchangeable step of preprocessing, since this paper does not focus on the method for building footprint extraction from the nDSM itself. In general, building footprints or other information on object geometries may originate from an arbitrary data source (e.g., from Volunteered Geographic Information).

For automated extraction of objects above ground from the DSM as conducted in this study, the first step is its normalization. For this purpose, a DTM (digital terrain model) is derived from the DSM, which contains height

information of the bare Earth's surface exclusively. Then, the DTM is subtracted from the DSM in order to calculate a normalized DSM (nDSM) which contains object heights only. The normalization process was carried out based on the methodology of morphological filtering presented in Geiß et al. (2015b).

Since objects above ground contain buildings as well as other objects that possess vertical extension (e.g., vegetation objects), the second step is segmentation and classification of the nDSM to extract building footprints that serve as multi-temporal image objects for the change detection workflow. Building footprint extraction was conducted according to Wurm et al. (2011), where a local contrast split segmentation algorithm is utilized for adaptive segmentation of the nDSM followed by classification of buildings by means of geometric features. This procedure is implemented as an iterative process in order to gradually separate individual buildings and to improve accuracy. The extracted building footprints are depicted in Fig. II-3. As mentioned above and also apparently visible from Fig. II-1, sensor viewing geometries (i.e., the off-nadir view angles of QuickBird and WorldView images) deviate significantly and cause severe object tilting. Besides viewing angles, the parallax distortion of buildings is additionally dependent on topography, the position within the image and object height. Since the DTM is flat within the experimental site and the areal extent of VHR imagery is small, topography and the position within the image can be neglected. However, this effect increases with increasing object height and therefore, this factor has to be considered especially in case of the high rise buildings present in the study area. In order to experimentally quantify the influence of deviating viewing geometries on the proposed unsupervised change detection approach, a simplistic but straightforward linear adjustment of building footprint geometries was applied. The adjustment was realized by systematic height-dependent shifting of building geometries separately in each image. For this purpose, building parallax distortion is measured in both images manually and linearly related to building height (mean height within each building footprint) from the nDSM. This experimental adjustment ensures that the same part of each building (i.e., the roof) is covered in both images and makes the content of image objects comparable (Fig. II-3). Accordingly, two sets of object geometries are available: original (for automated change detection; depicted in green color in Fig. II-3) and adjusted object geometries (for experimental quantification of the influence of deviating viewing geometries; displayed in blue in case of QuickBird data and orange for WorldView data in Fig. II-3b, respectively).

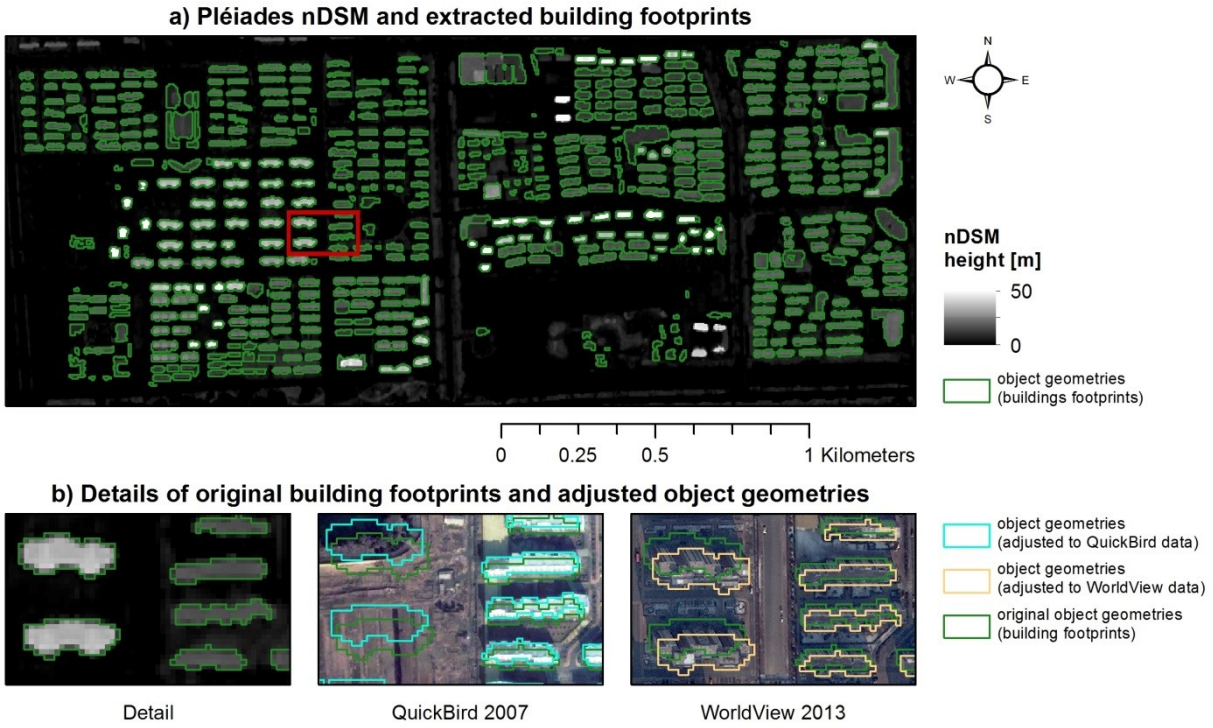

Fig. II-3 Object geometries and experimental spatial adjustment: a) Extracted building footprints and Pléiades nDSM, b) Detailed views of building footprints and adjusted object geometries according to QuickBird (t_0-1) and WorldView (t_0), respectively.

3.2 Preparation of object features

Based on the derived object geometries corresponding to building footprints and the preprocessed VHR imagery, 45 object-based features in the spectral and textural domains are extracted in order to capture comprehensive input information for change detection (Tab. II-1). *First*, common first order statistics of the multispectral data are utilized. *Second*, ratio features are included by means of simple ratios, normalized green, red and NIR (Sripada et al. 2006) as well as the normalized difference vegetation index (NDVI) and the normalized difference soil index (SAVI). *Third*, textural features are considered for the grayscale (calculated as mean intensity of blue, green and red) and NIR image channels. This distinction is incorporated in view of potential utilization of VHR imagery from aerial cameras which is commonly available as RGB data. For calculation of features based on the gray-level co-occurrence matrix (GLCM) (Haralick et al. 1973) as well as gray-level difference vector (GLDV) (Weszka et al. 1976), the grayscale as well as the NIR data sets are classified by means of the Jenks natural breaks algorithm (Jenks and Caspall 1971) into a reasonable quantization level of 5bit (32 gray levels) (Clausi 2002). Since the object geometries are defined constant over

time (i.e., multi-temporal objects), size and shape features are irrelevant in this setting. In the change detection analysis, the four feature sets are evaluated individually and additionally as a merged feature set including all 45 features. In addition, two sets of object geometries are utilized in the experiments: original (identical for both images) and experimentally adjusted (uniquely aligned to QuickBird and WorldView data, respectively) geometries.

Tab. II-1 Object-based feature sets

Multispectral Features (10)	Ratio Features (11)	Grayscale Texture Features (12)	NIR Texture Features (12)
mean intensity of blue channel	blue / green	GLCM angular 2nd moment	GLCM angular 2nd moment
mean intensity of green channel	blue / red	GLCM contrast	GLCM contrast
mean intensity of red channel	blue / NIR	GLCM correlation	GLCM correlation
mean intensity of NIR channel	green / red	GLCM dissimilarity	GLCM dissimilarity
mean intensity of all channels (brightness)	green / NIR	GLCM entropy	GLCM entropy
standard deviation of blue channel	red / NIR	GLCM homogeneity	GLCM homogeneity
standard deviation of green channel	normalized green[a]	GLCM mean	GLCM mean
standard deviation of red channel	normalized red[b]	GLCM standard deviation	GLCM standard deviation
standard deviation of NIR channel	normalized NIR[c]	GLDV angular 2nd moment	GLDV angular 2nd moment
maximum difference between channels	NDVI[d]	GLDV contrast	GLDV contrast
	SAVI[e]	GLDV entropy	GLDV entropy
		GLDV mean	GLDV mean

[a] normalized green = green / (NIR + red + green), [b] normalized red = red / (NIR + red + green), [c] normalized NIR = NIR / (NIR + red + green)
[d] NDVI = (NIR – red) / (NIR + red), [e] SAVI = [(NIR – red) / (NIR + red + L)] * (1 + L), where L = 0.5

The temporal relationship of the image objects in both images is included in terms of object-based image differencing of each feature. This technique has proven its effectiveness for change detection in VHR remote sensing (Bovolo and Bruzzone 2007b; Volpi et al. 2013) and reduces the information content to differences between the two acquisitions. According to this procedure, each feature holds information on changes (i.e., changed as well as unchanged objects) between the images. Due to differing characteristics of the features and thus, diverse ranges of values, normalization of the difference features is mandatory. Furthermore, clustering algorithms like k-means that are based on distance measures require normalized data for best results (Kantardzic 2011). According to Milligan and Cooper (1988), a suitable procedure is minimum-maximum normalization, which can be adapted to fit the desired value range of [-1, 1] as shown in (1).

$$v'(i) = 2 \cdot \left(\frac{v(i) - \min[v(i)]}{\max[v(i)] - \min[v(i)]} \right) - 1 \quad (1)$$

where $v'(i)$ is the normalized value of object i, $v(i)$ is the original value and $\min[v(i)]$ and $\max[v(i)]$ are the minimum and maximum values of all objects, respectively. This method of feature normalization preserves the variance of the input data and performs rescaling to a common value range of [-1, 1].

As the last step of feature preparation, the dimensionality of the data is reduced since most of the object features described in Tab. II-1 are highly correlated (e.g., first order statistics of the multispectral image channels) or contain redundant information such as some GLCM and GLDV texture measures (e.g., GLCM Homogeneity and Dissimilarity). In general, clustering high dimensional data is a challenging task since the likelihood of presence and number of irrelevant features grows with dimension (Berkhin 2006). In addition, dimensionality reduction is always desirable in order to reduce time and space complexity for computing. Generally, two different approaches for dimensionality reduction exist: the *selection* of most valuable *features*, termed *feature selection*, and the *transformation* of *data*, termed *feature extraction* (Webb 2011). Since feature selection procedures mostly require labeled data or "involve a trial-and-error process" (Jain et al. 1999) related to the complete clustering workflow, feature extraction is preferred in the presented unsupervised change detection context.

A common and frequently applied technique for linear feature extraction is principal component analysis (PCA). PCA seeks for uncorrelated linear combinations of the original data which are sorted in decreasing order of importance. The importance of each principal component (PC) is expressed by its corresponding eigenvalue of the correlation matrix, which is directly related to the proportion of explained variance with respect to the complete data set. In this study, the normalized differences of features are used as input for PCA.

3.3 Selection of relevant principal components

A crucial step in PCA is the selection of the number of relevant (i.e., non-trivial) components. This represents the trade-off between preservation of information and dimensionality reduction. Several stopping rules for PCA are presented in Jackson (1993), Ferré (1995) and Peres-Neto et al. (2005), whereas all of these techniques are strongly influenced by the nature of the data and thus, determination of relevant PCs requires consideration of at

least more than one method. In the current study, three methods are combined uniquely in a multi-objective manner.

First Bartlett's test for the first PC is applied in order to ensure that its eigenvalue is significantly different from the remaining eigenvalues and thus, at least one non-trivial component is contained in the data (Jackson 1993). The test is designed as a statistical hypothesis test based on the χ^2 distribution with $p(p-1)/2$ degrees of freedom according to (2).

$$\chi^2 = -\left[n - \frac{1}{6}(2p+11)\right] \ln |R| \qquad (2)$$

where $|R|$ is the determinant of the correlation matrix, n is the number of observations and p is the number of features in the original data. If this test is significant and at least the first PC entails meaningful information, a *second* condition is that the cumulative proportion of variance (Ferré 1995) explained by the considered PCs exceeds a certain threshold. This simple technique is commonly used, but it should be considered as a descriptive indicator only (Ferré 1995). In the current study a widely-used critical value of 0.9 is chosen (Ferré 1995). *Third*, in order to add a more objective and substantial rule for selection of the number of relevant components, an averaged parallel analysis as proposed by Peres-Neto et al. (2005) is applied. This technique generates random, uncorrelated data in the dimension of the original data and conducts a PCA in order to retrieve the eigenvalues for all PCs. This procedure is repeated 1,000 times and the resulting eigenvalues of the random data are averaged. The resulting randomized eigenvalues are compared to those of the original data and the number of relevant components is determined according to the point where the original eigenvalues exceed the randomized eigenvalues. In this study, the technique by Peres-Neto et al. (2005) is modified in a way that even one PC with lower eigenvalue compared to the corresponding randomized eigenvalue is included in the relevant set of PCs, i.e., the utilization in this approach is less rigorous to avoid unintended loss of relevant information on changes. The final number of non-trivial PCs is determined according to the greater number of retained components from both techniques (proportion of variance threshold and averaged parallel analysis) to basically consider both methods and to ensure a suitable trade-off between preservation of relevant information and reduction of dimensionality.

The relevant PCs of each feature set (multispectral, ratios, grayscale texture, NIR texture and merged data set) provide the input for the subsequent cluster analysis for unsupervised change detection.

3.4 Cluster analysis for change detection

Clustering is an exploratory data analysis technique that aims at finding structural groups present in the data (Webb 2011). Although it does not provide category labels, clustering has proven to be well suited for unsupervised change detection in remote sensing (Celik 2009; Ding et al. 2015; Ghosh et al. 2011; Volpi et al. 2010, 2012; Zheng et al. 2014b). In the current study, semantic information on the desired classes (i.e., changed and unchanged buildings) is given implicitly by appropriate preparation of the input features (see Section 0). At the top-most level, the vast number of clustering algorithms available in literature (Xu and Wunsch 2010) can be divided into hierarchical and partitioning methods. In general, partitioning methods are preferred in pattern recognition due to the nature of the data (Jain 2010), i.e., due to high dimensionality and/or large number of data points (Gan et al. 2007). In the current study, the widely used partitioning clustering algorithm k-means (Hartigan and Wong 1979) is utilized. The k-means algorithm has proven its effectiveness – although in a pixel-based manner – in several studies of change detection in remote sensing (Celik 2009; Zheng et al. 2014b). The number of clusters is determined by the nature of the two-class classification problem in this particular application (i.e., identification of changed and unchanged objects) and is thus set to "two" *a priori*. Notably, this avoids the challenging problem to estimate a suitable number of clusters in an automated way, e.g., by means of comparative analysis of cluster validity measures (Halkidi et al. 2001).

The k-means algorithm, which was initially proposed almost 50 years ago, is still widely used today (Berkhin 2006; Jain 2010). The key advantages of k-means are its conceptual simplicity, easy implementation, efficiency and its empirical success (Jain 2010). K-means relies on a distance measure for quantification of (dis)similarity of data points, commonly the Euclidian distance in case of numeric data (Jain 2010). Further input parameters are the number of desired clusters and an initialization of cluster centers, which is mostly chosen randomly (Gan et al. 2007; Hartigan and Wong 1979). With these user-specific input parameters, k-means partitions the data iteratively as follows (Jain 2010):

1. Calculate the (Euclidian) distances between all data points and the centers of the desired clusters
2. Assign each data point to the closest cluster center
3. Compute refined cluster centers using the mean value of all associated data points
4. Repeat steps 1-3 until cluster membership stabilizes, i.e., the sum of squared distances for all data points associated to the cluster centers becomes minimal

Concluding, k-means is based on a minimization of a (Euclidian) distance-based square-error objective function (Kantardzic 2011). It creates spherical ("ball-shaped") clusters in the input data space, divided by linear separating hyperplanes (Jain 2010). In this study, the implementation presented by Hartigan and Wong (1979) is used which includes several random initializations of cluster centers. Finally, a single manual intervention is required to assign the correct class labels to the two resulting clusters in order to enable accuracy assessment.

3.5 Accuracy assessment

For performance evaluation of the labeled clustering result on different feature sets, accuracy is assessed in terms of the Cohen's kappa coefficient κ (Congalton 1991). A key advantage of the kappa statistic is that it incorporates the off-diagonal elements of the error matrix, i.e., omission and commission errors are included (Congalton 1991). The value of κ shows the randomness of the results, i.e., $\kappa = 0$ denotes complete statistical randomness while $\kappa = 1$ and $\kappa = -1$ indicate perfect positive and negative statistical connectivity, respectively. Reference data were collected by simultaneous visual inspection of the QuickBird and WorldView VHR remote sensing imagery area-wide across the experimental site comprising changed and unchanged labels for all 564 buildings (215 changed buildings, 349 unchanged buildings). This data set will be used for all validation purposes and is referred to as the reference map (for a visual illustration see maps in Fig. II-7 and Fig. II-8, respectively).

4. Experimental results

The following results of the applied methodology are presented for both, the original objects and the spatially adjusted object geometries as shown in Fig. II-3. To avoid redundancies, the evaluation based on the original object geometries is emphasized and the comparison is drawn to the adjusted geometries where applicable in order to demonstrate the influence of deviating viewing geometries. The presented approach was implemented under completely unsupervised conditions.

4.1 Selection of relevant principal components

After calculation of the difference features (Tab. II-1) together with normalization and subsequent PCA analysis, selection of relevant PCs was a crucial step in data preparation. PCA was utilized in order to transform correlations and redundant information of the input features to PCs of higher order, while meaningful information for change detection is concentrated in the first few PCs (also see Fig. II-6) and thus, a reasonable cut-off point for the relevant PCs must be determined. *First,* Bartlett's test for the first PC was conducted in order to ensure that it significantly differs from the remaining PCs and thus, at least one non-trivial PC is contained in the data. The test was highly significant for all feature sets according to p-values, i.e., values of $p<0.001$ were reached throughout. Thus, it was ensured that the first PC was relevant (i.e., at least the first PC can be retained) for all feature sets.

Secondly, cumulative proportion of explained variance of the PCs was evaluated. Fig. II-4 depicts the graphs for the four individual feature sets as well as the merged set of all 45 features. The proportion of explained variance of the individual feature sets (Fig. II-4a-d) exceeded the threshold value of 0.9 mostly after inclusion of two, at maximum three (in case of NIR texture features) PCs. Saturation effects (cumulative variance contained in remaining PCs < 0.01) occurred after around four PCs. For the merged data set, five PCs were required in order to achieve sufficient proportion of explained variance. Reasonably, with higher total number of PCs, the proportion of explained variance accumulates less rapidly and saturation effects do not occur until 12 PCs (Fig. II-4e).

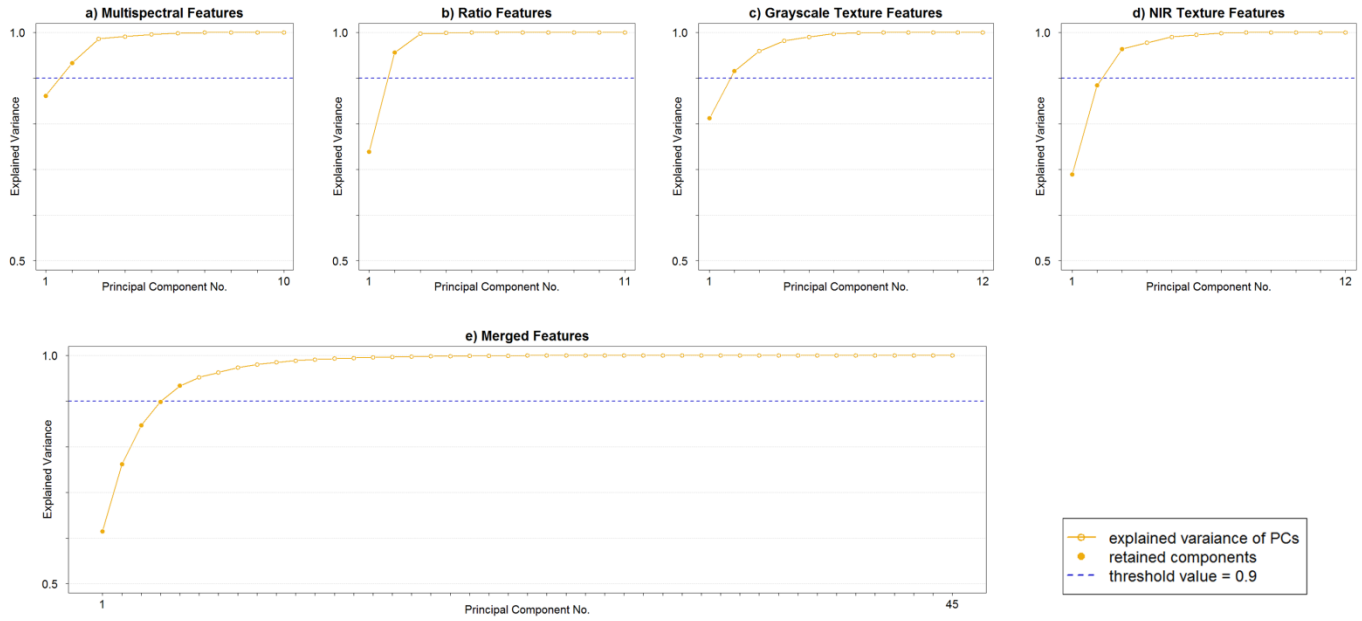

Fig. II-4 Cumulative proportion of explained variance in case of original object geometries after PCA: a) Multispectral features b) Ratio features c) Grayscale texture features d) NIR texture features e) Merged features.

The *third test* for selection of relevant PCs was an averaged parallel analysis approach. Fig. II-5 depicts the eigenvalues of the PCs of each feature set against their randomized counterparts, whereas the point of intersection indicates the separation of meaningful data against random information. For the individual feature sets, two (Fig. II-5a) and three (Fig. II-5b-d) PCs were retained respectively, whereas in case of the merged feature set six PCs were included according to this analysis. It can be observed that the absolute eigenvalue of the first PC of the merged feature set is substantially higher (approximately 28) compared to those of the individual feature sets (<10). This indicates that features among different sets are correlated (e.g., texture features based on grayscale and NIR data) and consequently, they can be subsumed in the first PC of the merged data set.

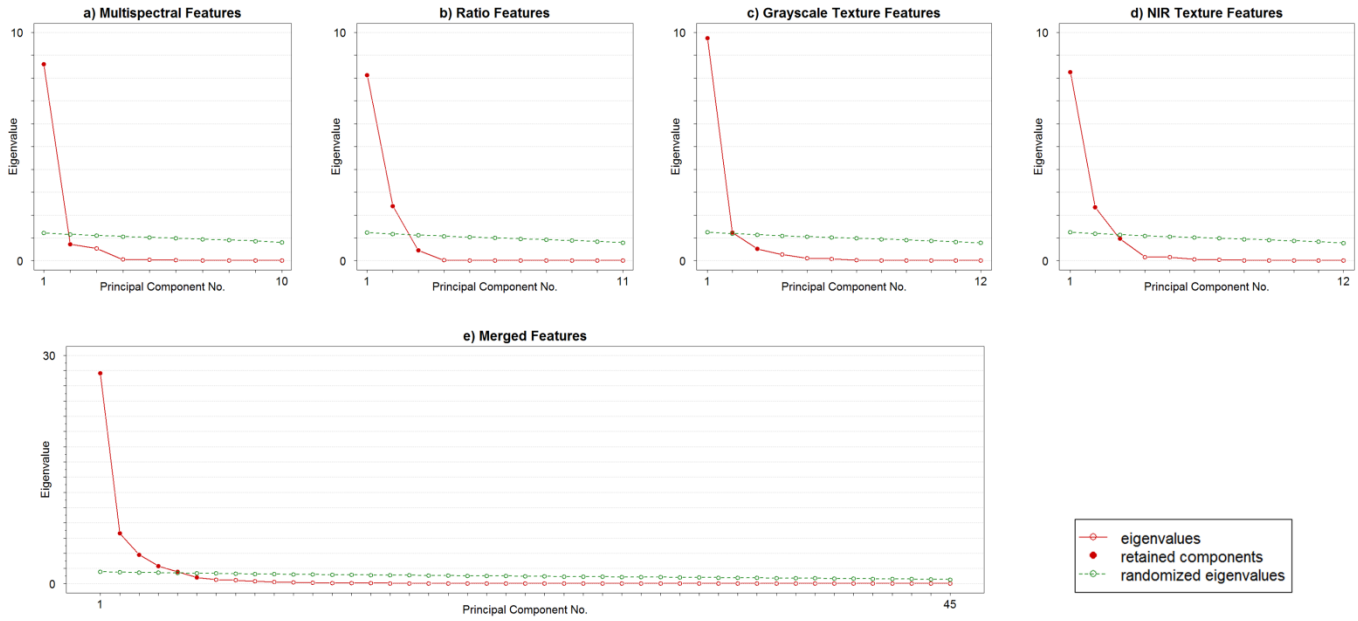

Fig. II-5 Averaged parallel analysis in case of original object geometries after PCA: a) Multispectral features b) Ratio features c) Grayscale texture features d) NIR texture features e) Merged features.

As explained in Section 0, the number of relevant PCs was determined according to higher number of retained PCs in both analyses in order to consider both methods and to preserve relevant information with simultaneous reduction of dimensionality. Results of PCA analysis for both data settings, i.e., original and adjusted object geometries, led to very similar results and thus, the selected number of PCs was identical (Tab. II-2).

Tab. II-2 Number of relevant Principal Components

	Original Object Geometries	Adjusted Object Geometries
Multispectral features	2	2
Ratio features	3	3
Grayscale texture features	3	3
NIR texture features	3	3
Merged features	6	6

4.2 Clustering for change detection and accuracy assessment

The statistical distributions of the PCs according to the five groups of features are visualized in Fig. II-6 to demonstrate the distribution of the data subsequent to PCA transformation in relation to class labels (unchanged and changed objects according to reference data).

Each reduced set of PCs contains substantial proportion of explained variance (Fig. II-6, also see Fig. II-4) with respect to the two target classes and thus, relevant information for change detection in the experimental site. Furthermore, the general characteristic in PCA of ordering the PCs in decreasing importance according to the input information is retained by the proposed approach. Concluding, the most important information with respect to changes (i.e., discriminative power according to changed and unchanged image objects) is stored in the first few PCs, while PCs of higher order tend to contain less important information, like correlations and redundancies between the original input features. The reduced set of PCs serves as input for subsequent k-means clustering.

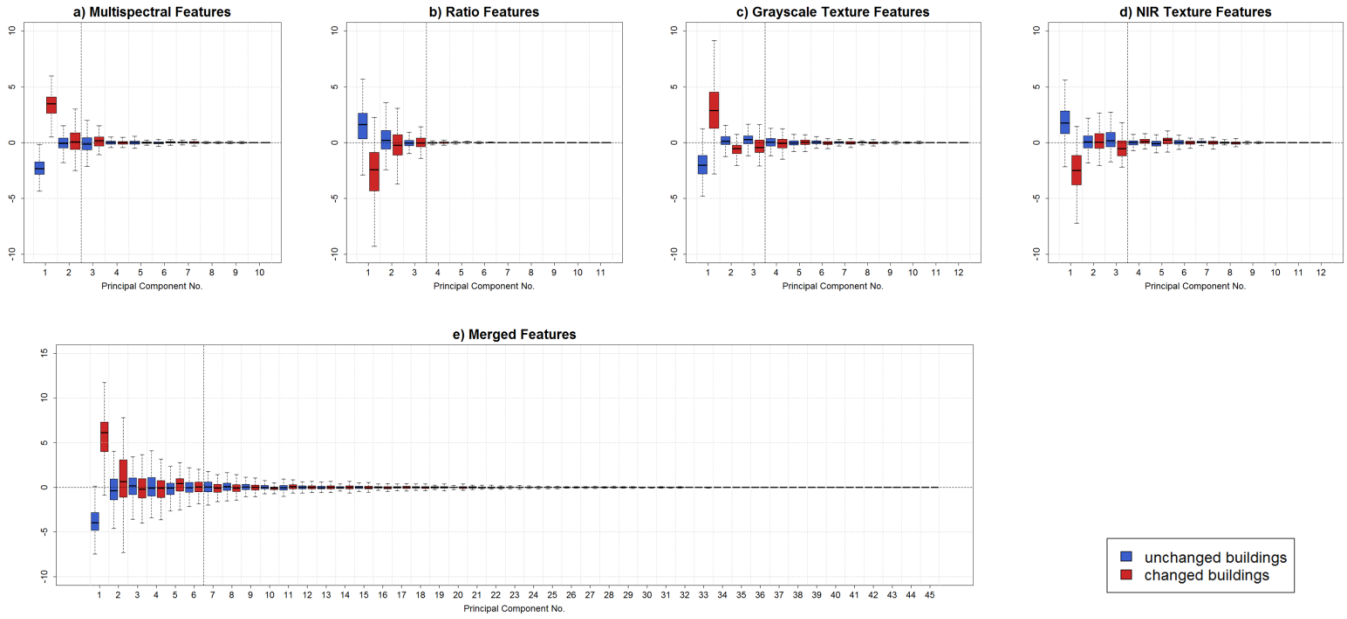

Fig. II-6 Boxplots of PCs with class labels (unchanged buildings in blue, changed buildings in red) from reference data in case of original object geometries: a) Multispectral features b) Ratio features c) Grayscale texture features d) NIR texture features e) Merged features. The vertical black dashed lines indicate the cut-off point for the relevant PCs.

Fig. II-7 displays the change detection results in terms of change maps and corresponding PCA feature spaces (Fig. II-7b-e only visualizes the first and second dimension of the three- resp. six-dimensional feature space). Tab. II-3 holds the associated measure of accuracy with respect to kappa statistics (for original object geometries as depicted in Fig. II-7 see column 1 of Tab. II-3).

Best results were achieved based on the multispectral ($\kappa = 0.87$) and merged ($\kappa = 0.88$) feature sets (Fig. II-7a and Fig. II-7e). The scatterplots of these feature sets possess a distinct bimodal distribution (i.e., clear separation) accounting for the two desired classes in the two-dimensional feature space (green circles in Fig. II-7a and Fig. II-7e). Ratio features led to worst accuracies in terms of Kappa of 0.73. Within the ratio feature space, the desired classes were overlapping (red circle in Fig. II-7b) and thus, it was challenging for k-means clustering to separate them. This becomes visible in the eastern part of the corresponding change map of Fig. II-7b, where separation of changed and unchanged buildings is deficient. Less severe overlapping of classes also accounts for the grayscale and NIR texture feature sets and complicates correct classification of changed and unchanged objects (Fig. II-7c and Fig. II-7d). Another important characteristic of the grayscale and NIR texture features in the PCA input feature space were outlying image objects in the PCA feature space (red circles in Fig. II-7c and Fig. II-7d) which correspond to the subordinate class of changed buildings that originate from water bodies at t_0-1 (also see Fig. II-1). Due to the general characteristics of texture, these areas appear very homogeneous at t_0-1 (i.e., water bodies), whereas at t_0 these features possess more heterogeneous textural characteristics of buildings. Nevertheless, these outlying data points were classified correctly in both feature sets. The outliers originating from the texture feature sets can be traced in the merged feature set (Fig. II-7e, red circle in the lower right corner of the scatterplots), whereas the mixture of all 45 features mitigates the influence of outliers on the overall distribution of the data and thus, negative influence on the classification result and corresponding accuracy is anticipated implicitly.

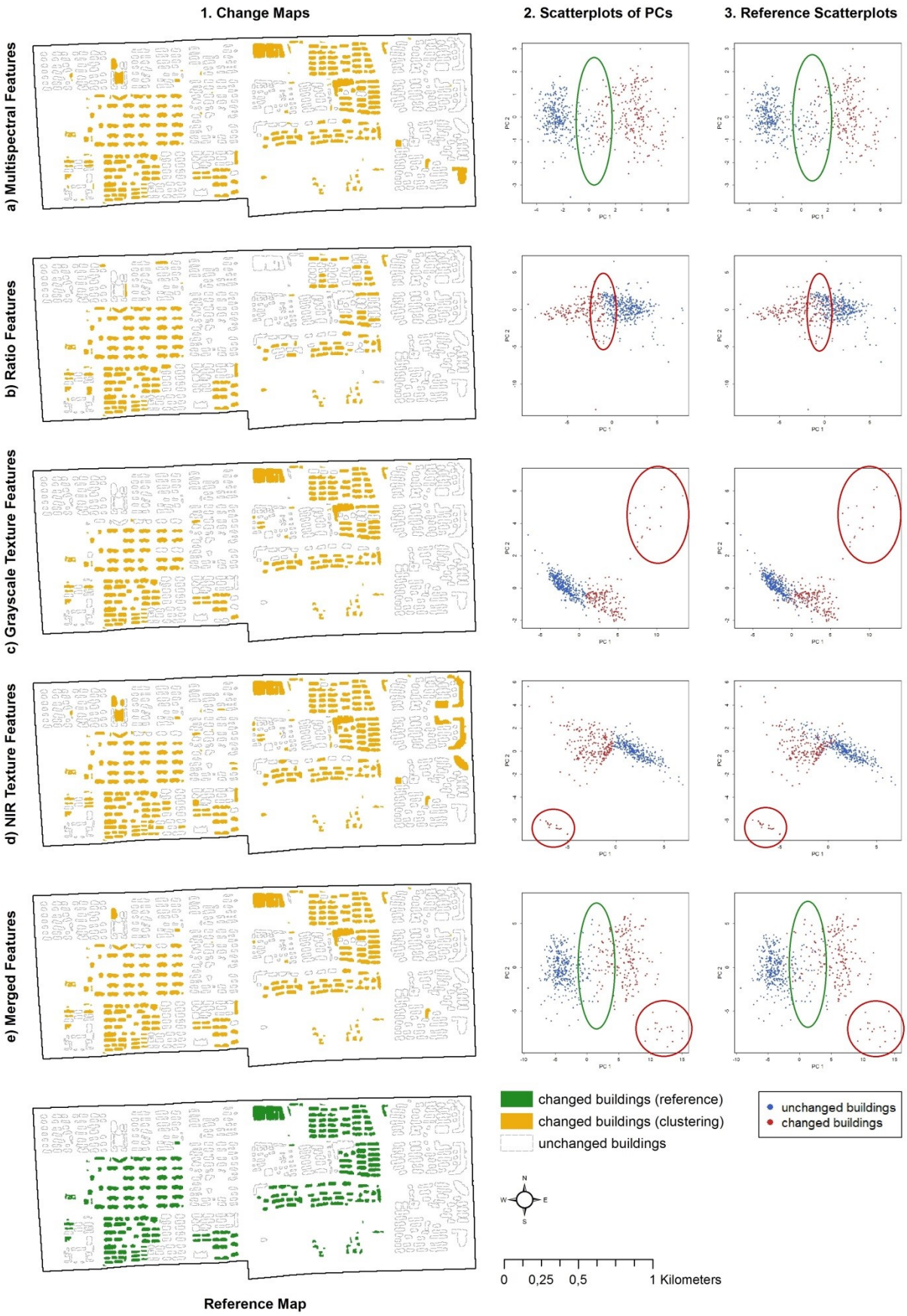

Fig. II-7 Results of change detection in case of original object geometries: a) Multispectral features b) Ratio features c) Grayscale texture features d) NIR texture features e) Merged features. Column 1 shows

change maps and reference map (bottom), column 2 depicts the classification result as scatterplots of PC 1 (horizontal axis) against PC 2 (vertical axis), column 3 displays scatterplots of PC1 against PC2 colored according to reference classification. Green circles indicate good classification performance, red circles refer to worse classification agreement, respectively.

Tab. II-3 Accuracy (κ) according to original and adjusted object geometries of the proposed approach

	Original Object Geometries	Adjusted Object Geometries
Multispectral features	0.87	0.90
Ratio features	0.73	0.73
Grayscale texture features	0.81	0.84
NIR texture features	0.78	0.78
Merged features	0.88	0.92
Mean of feature sets	**0.81**	**0.83**

The influence of deviating viewing geometries is estimated according to differences of accuracy in terms of the Kappa statistics (Tab. II-3) and is displayed exemplarily in case of the merged feature set in Fig. II-8. Generally, the experimental adjustment procedure enhanced κ in the order of 0.03 to 0.04 or at least led to identical accuracy in case of the ratio and NIR texture feature sets (Tab. II-3). It can be deduced from the comparison of scatterplots of original and adjusted object geometries (Fig. II-8a and Fig. II-8b), that the adjustment procedure induced slight concentration of data points in the PCA feature space. Overall, the merged feature set returned the best change detection results in both cases of original ($\kappa = 0.88$) and adjusted ($\kappa = 0.92$) object geometries (Fig. II-8).

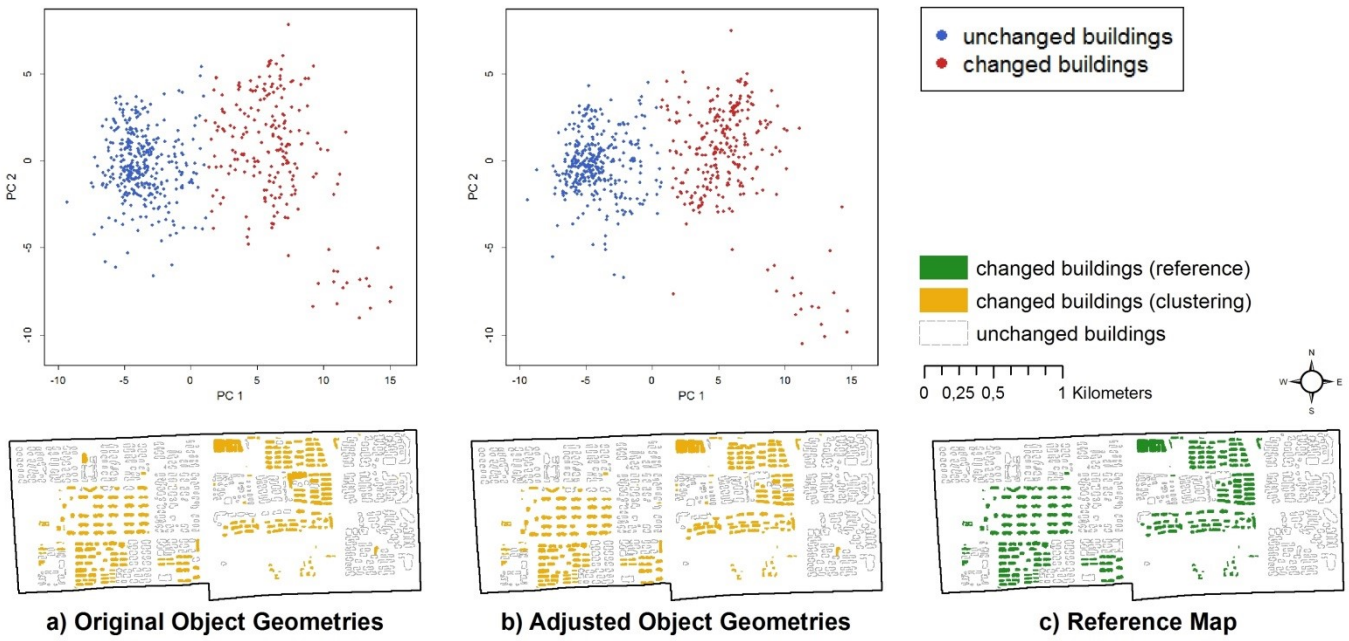

Fig. II-8 Comparison of results of the unsupervised change detection approach based on the merged feature set: a) Original object geometries b) Adjusted object geometries (displayed geometries correspond to the WorldView image). Scatterplots (top) of PC 1 (horizontal axis) against PC 2 (vertical axis) and corresponding change maps (bottom). c) Reference map.

5. Discussion

This study presents an unsupervised approach for object-based change detection in urban environments with focus on individual buildings. In contrast, other studies that aim at identification of changes in urban areas based on VHR data often refer to general changes of land cover and do not focus at thematic objects of buildings (Im et al. 2008). Several studies that focus on individual buildings are related to the applications of damage detection of buildings based on multi-temporal VHR data (Al-Khudhairy et al. 2005; Chen and Hutchinson 2007) or updating of building inventory maps which only use a single VHR image together with object geometries (Bouziani et al. 2010; Matikainen et al. 2010). Nevertheless, some studies that are closely related to the approach presented in this paper exist in literature. Klonus et al. (2012) apply a Combined Edge Segment Texture (CEST) analysis to multi-temporal QuickBird imagery and achieve up to 80% overall accuracy. Tang et al. (2013) employ multi-temporal matching of a morphological building index in pairs of QuickBird and IKONOS images reaching values of quality in the order of 50 to 70%. A similar methodology is proposed by Huang et al. (2014) who obtain errors of less than 10% based on multi-temporal QuickBird data. In contrast to these studies, the presented

approach is capable of handling multi-temporal VHR data from different sensors achieving comparable values of accuracy (i.e., κ around 0.8 to 0.9). A recent study that employs data from different sensors is Wang et al. (2015), where QuickBird and WorldView data is used for land cover change detection in urban environments reaching values of κ around 0.5 compared to superior κ in the order of 0.8 to 0.9 in this study.

Due to the unsupervised implementation of algorithms as well as the capability of handling multi-temporal VHR data from different sensors, transferability is generally favored by the proposed approach. With respect to required input data for change detection, a multi-temporal pair of VHR imagery (possibly from different sensors) together with object geometries is needed. In this context, the object geometries can be delineated from a nDSM as presented in this study, but can equally originate from an arbitrary data source (e.g., from Volunteered Geographic Information). In relation to transferability to different cities, it can be assumed that the methodology is capable of detection of changes in any city worldwide based on the independent and objective nature of remote sensing data acquisition. Nevertheless, in case that one of the target classes possesses only small prior probability (i.e., the distribution of classes is strongly imbalanced), the transferability of the proposed approach must be investigated in future research. Such a setting may be caused by a small (e.g., urban growth in less dynamic cities) or very high (e.g., long-term observation of urban growth in highly dynamic cities) amount of changes in other urban areas. Another crucial point with respect to transferability may be the matching of object geometries and image content (i.e., pixels related to image objects), especially if the city structure is more complex (compared to the distinct structure of Dongying in this study) or distortion effects due to topography or other non-linear effects are present. As experiments revealed, object geometries may be adjusted for highest accuracies in case of area-wide consideration of deviating viewing geometries. Especially at large spatial scales, when topography or large-scale non-linear effects like the position of objects within the image become more influential (see Section 3.1.2), a more advanced and automatic method for (height-dependent) shifting of building geometries should be employed compared to the experimental method applied in this study. Nevertheless, the experiment presented in this study underlines the robustness of the approach since accuracy only improved very slightly and thus, the influence of deviating viewing geometries turned out to be less severe. In addition, the proposed workflow is designed for change detection within the predefined object geometries, i.e., in the current data setting (object geometries matching t_0) only an increase of building inventory (or major alterations of buildings) can be detected. This corresponds to the analysis of urban growth from a thematic point of view. Accordingly, in the opposite data setting when object geometries correspond to t_0-1, the thematic scope of the proposed approach can be reversed and demolition of buildings can be detected. This aspect will be

examined in future work in order to transfer the change detection approach to further thematic fields of application (e.g., damage detection) in the context of urban remote sensing.

The methodology of the proposed approach is based on PCA-based transformation of object-based difference features followed by k-means clustering for change detection. Although the approach returned viable and robust results, the most suitable representation of feature sets for clustering must be examined in future work. Besides linear PCA, also non-linear approaches such as kernel PCA (Schölkopf et al. 1998) may be evaluated in order to increase prediction accuracy. Another possibility for incorporation of nonlinearity is utilization of different clustering approaches that do not require additional input parameters for unsupervised classification of changes. In general, the choice of clustering algorithms is not limited to k-means and different approaches that basically satisfy the needs of an unsupervised change detection approach (i.e., that do not require additional input parameters besides the number of clusters) may be considered. To this point, the change detection approach is only capable of capturing the major type of change due to the basic characteristics of unsupervised implementation of a two-class classification procedure.

6. Conclusion

This paper presents a robust and flexible approach for unsupervised change detection of buildings which is capable of handling remote sensing data acquired by different sensors. Such a setting is very likely in practical applications, where suitable image pairs from the same sensor of a particular area of interest for a specific period of time are largely absent. Necessary object geometries were extracted from a nDSM but can equally originate from an arbitrary data source (e.g., from Volunteered Geographic Information). These aspects are beneficial regarding transferability to other VHR remote sensing data, different study areas or further fields of application. As another asset from a practical point of view, the workflow is implemented in a completely unsupervised way in order to ensure flexibility and transferability of the proposed approach and to enable highly automated systems for processing of vast amounts of VHR remote sensing data. In detail, a tripartite procedure for selection of relevant principal components was employed as part of data preprocessing. Changed and unchanged buildings were finally discriminated using k-means clustering whereas in addition, the performance of different feature sets was assessed. Generally, the proposed approach turned out to be well-suited for unsupervised two-class change detection. Best overall results were achieved based on a merged feature set which incorporated ratio features, first order statistics of the multispectral image channels as well as texture features deduced from the

grayscale and NIR image. A reasonable alternative is the exclusive use of multispectral features which performed almost equivalently well. Since grayscale texture features outperformed NIR texture features, utilization of RGB data (e.g., from aerial cameras) can be an eligible alternative to higher-level RGB+NIR data. Deviating viewing geometries, which occur likely in practical applications (e.g., when optical VHR images from different satellite systems are employed), were found to deteriorate the change detection result in the order of up to 0.04 according to κ. Hence, this experiment underlines the robustness of the proposed approach since accuracy improved only slightly with correction of object geometries and thus, the influence of object tilting was found less severe in the presented data setting.

Acknowledgements

This research was conducted in the context of the DELIGHT project (www.delight.eoc.dlr.de) funded by the German Federal Ministry of Education and Research, BMBF. Funding codes 02WCL1249A and 02WCL1249I. The authors would like to thank Matthias Boes (European Space Imaging, EUSI) for providing the WorldView imagery and Dr. Klaus Martin (Company for Remote Sensing and Environmental Research, SLU) for his support.

Chapter III. Evaluation of clustering algorithms for unsupervised change detection in VHR remote sensing imagery

Proceedings of the 2017 Joint Urban Remote Sensing Event (JURSE), IEEE Xplore (2017), 1-4.

Tobias Leichtle[a, b, c], Christian Geiß[b], Michael Wurm[b], Tobia Lakes[c], and Hannes Taubenböck[b]

[a] Company for Remote Sensing and Environmental Research (SLU), Kohlsteiner Straße 5, 81243 München, Germany

[b] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchner Straße 20, 82234 Weßling, Germany

[c] Humboldt-Universität zu Berlin, Geography Department, Rudower Chaussee 16, 12489 Berlin, Germany

https://doi.org/10.1109/JURSE.2017.7924625
Date of Conference: 06th March - 08th March 2017, Dubai, UAE; Date Added to IEEE Xplore: 11th May 2017

Abstract

Remote sensing has proven to be an adequate tool for observation of changes to the Earth's surface. Especially modern space-borne sensors with very-high spatial resolution offer new capabilities for monitoring of dynamic urban environments. In this context, clustering is a well suited technique for unsupervised and thus highly automatic detection of changes. In this study, seven partitioning clustering algorithms from different methodological categories are evaluated regarding their suitability for unsupervised change detection. In addition, object-based feature sets of different characteristics are included in the analysis assessing their discriminative power for classification of changed against unchanged buildings. In general, the most important property of favorable algorithms is that they do not require additional arbitrary input parameters except the number of clusters. Best results were achieved based on the clustering algorithms k-means, partitioning around medoids, genetic k-means and self-organizing map clustering with accuracies in terms of κ statistics of 0.8 to 0.9 and beyond.

1. Introduction

Monitoring of changes is one of the intrinsic capabilities of remote sensing. With respect to increasingly available remote sensing images with very-high spatial resolution (VHR), highly automatic techniques for detection of changes become more and more important (Bruzzone and Bovolo 2013). Especially in urban environments, where different surface materials and image objects are concentrated in a spatially complex and highly dynamic manner, robust and simultaneously flexible change detection is a challenging task. Particularly the high level of detail in VHR images within urban environments favors object-based methods over rather traditional pixel-based methods for automatic detection of changes (Hussain et al. 2013). Regarding rapidly changing urban areas in emerging and developing countries of Africa and Asia, monitoring of urban growth in terms of expanding building inventory is a common task (Taubenböck et al. 2012). For these reasons, an unsupervised object-based change detection approach with focus on the high detail of individual buildings is intended in this study.

In this context, a promising and recently emerging methodology for unsupervised change detection is partitioning clustering (Jain 2010). For example, Celik (2009) and Zheng et al. (2014b) apply k-means clustering

to difference images of optical and SAR imagery, respectively. Kernel k-means is used for change detection in QuickBird images by Volpi et al. (2010), while the same group of authors extended their approach and applied it to SPOT and Landsat data in Volpi et al. (2012). Ghosh et al. (2011) employ two fuzzy clustering algorithms for binary change detection in Landsat difference images. A technique from a different clustering domain - sparse hierarchical clustering - is used by Ding et al. (2015) for change detection in VHR imagery. It has to be noted that all of the mentioned studies either work directly on pixel-level or incorporate the spatial domain on a local neighborhood level (i.e., directly neighboring pixels) only. Furthermore, only a small fraction of these studies employs VHR imagery.

The objective of this study is to evaluate and discuss the capability of different partitioning clustering algorithms for unsupervised two-class change detection of individual buildings using VHR imagery. Seven clustering algorithms that basically satisfy the needs of unsupervised change detection are evaluated since the most qualified algorithm for any specific application cannot be determined a priori. Furthermore, comprehensive feature sets of different characteristics are included in the evaluation in order to indicate the basic suitability of distinct types of features for change detection of buildings regarding different algorithms.

2. Study area and data sets

The evaluation of clustering algorithms for unsupervised change detection is conducted in the city of Dongying, China, which is located in the Yellow River Delta south of Beijing. Dongying is a typical example of emerging Chinese cities and exhibits dynamic urban growth. The experimental site within the study area covers a characteristic excerpt of the city comprising about 550 buildings.

Building geometries are available from segmentation of a normalized Digital Surface Model (nDSM) with 4m spatial resolution (Wurm et al. 2011), which was acquired on October 18, 2013 (t_0). The building geometries correspond to building footprints due to their derivation from the nDSM. Optical imagery with very-high spatial resolution is available by means of QuickBird data acquired on February 25, 2007 (t_0-1) and WorldView data recorded on January 17, 2013 (t_0), respectively. Both image products possess a spatial resolution of 50cm and are available with four spectral bands (blue, green, red, NIR). Although the individual bands of the two sensors cover slightly different wavelengths, the deviation of their respective mean values is marginal. The optical images were geometrically adjusted with a root mean square error (RMSE) of co-registration of 1.5 pixels (i.e., 75cm). In addition, the imagery was atmospherically corrected using the Atmospheric and Topographic

Correction (ATCOR) toolbox (Richter 1996) and radiometrically adjusted by means of band by band histogram matching.

3. Methodology

The workflow of this study is structured according to Leichtle et al. (2017a) and conducted as follows: First, different sets of object-based features are calculated based on multi-temporal VHR imagery and building geometries. Second, change detection in terms of unsupervised discrimination of changed and unchanged buildings is implemented by means of clustering. Finally, seven partitioning clustering algorithms are opposed and validated for evaluation of their applicability for change detection in the context of urban growth at the building level.

3.1 Generation of object-based features

Different sets of object-based features (Tab. III-1) are calculated based on each VHR data set (i.e., QuickBird at (t_0-1) and WorldView at t_0) and common object geometries (i.e., building footprints). The multi-temporal features of each building are subtracted in terms of t_0 - (t_0-1) in order to directly incorporate the temporal relationship (i.e., information on possible changes between the two images). The employed features comprise a comprehensive selection as listed in Tab. III-1. In addition to the three feature sets of multispectral (MS), ratio and grayscale texture features, a merged feature set consisting of all 33 features is considered.

Tab. III-1 List of object-based features

Multispectral (MS) Features (10)	Ratio Features (11)	Grayscale Texture Features (12)
mean intensity blue band	blue / green	GLCM angular 2nd moment
mean intensity green band	blue / red	GLCM contrast
mean intensity red band	blue / NIR	GLCM correlation
mean intensity NIR band	green / red	GLCM dissimilarity
mean intensity all bands	green / NIR	GLCM entropy
standard deviation blue band	red / NIR	GLCM homogeneity
standard deviation green band	normalized green[a]	GLCM mean
standard deviation red band	normalized red[b]	GLCM standard deviation
standard deviation NIR band	normalized NIR[c]	GLDV angular 2nd moment
maximum difference between bands	NDVI[d]	GLDV contrast
	SAVI[e]	GLDV entropy
		GLDV mean

grayscale image = (blue + green + red) / 3

[a] normalized green = green / (NIR + red + green), [b] normalized red = red / (NIR + red + green), [c] normalized NIR = NIR / (NIR + red + green)

[d] NDVI = (NIR – red) / (NIR + red), [e] SAVI = [(NIR – red) / (NIR + red + L)] * (1 + L), where L = 0.5

3.2 Change detection workflow

First, the object-based features were normalized using minimum-maximum normalization to a common range of values of [-1, 1]. Second, a principal component analysis (PCA) is conducted in order to eliminate redundancies that are present in the comprehensive input data set. Since PCA rearranges the data by transformation of redundant information to principal components (PCs) of higher order, the number of relevant PCs is determined applying a tripartite procedure as presented in Leichtle et al. (2017a).

The initial analysis for selection of relevant PCs is a Bartlett's test, which determines if the first PC significantly deviates from the remaining PCs (i.e., if structure with respect to the PCs is present in the data). This is followed by investigation of cumulative proportion of variance, whereas the summarized variance of the considered PCs must exceed the common threshold of 0.9. Subsequently, an averaged parallel analysis is conducted, which returns the number of PCs associated to their respective eigenvalue exceeding the eigenvalue of random data of equal dimensions. The second and third conditions are wrapped up by means of a maximum rule in order to consider both methods. More details on the selection procedure for relevant PCs can be found in Leichtle et al. (2017a).

Finally, the relevant PCs are classified into the two target groups (i.e., changed and unchanged buildings) by means of clustering. For this purpose, partitioning clustering algorithms are employed and evaluated in this study.

3.3 Evaluation of clustering algorithms

Seven relevant and suitable partitioning clustering algorithms from different methodological categories are evaluated since the most qualified algorithm cannot be chosen a priori (Jain 2010). With respect to the scope of unsupervised change detection, the prerequisite to relevant techniques is the potential of fully automatic implementation, i.e., no additional arbitrary input parameters for processing besides the number of clusters are required (or otherwise a reliable technique for the estimation of these parameters is available). In addition, the algorithms under study are intended to cover different methodological categories and represent well-established approaches in partitioning clustering. Thus, the following techniques are chosen in this study (with utilized R packages):

1. k-means [stats]
2. partitioning around medoids (PAM) [cluster]
3. genetic k-means [skmeans]
4. kernel k-means [kernlab]
5. spectral clustering [kernlab]
6. self-organizing map (SOM) clustering [som]
7. expectation-maximization (EM) clustering [mclust]

Most clustering algorithms under study can be implemented and initialized in a completely unsupervised way when the number of clusters is set to “two” (referring to changed and unchanged buildings, respectively) a priori. With respect to arbitrary input parameters, the EM clustering algorithm encompasses selection of an appropriate mixture model, which can be initialized using a Bayesian model selection procedure as shown in Fraley and Raftery (2002). For kernel k-means and spectral clustering, the Gaussian RBF kernel is utilized due to its versatility (Karatzoglou et al. 2004). The Gaussian RBF kernel transformation comprises the arbitrary parameter of inverse kernel width within the kernel function (Karatzoglou et al. 2004). This parameter must be initialized and is estimated by a common heuristic using the median value within the 0.1 and 0.9 quantile of the mean squared error of the input data (Karatzoglou et al. 2004). Since the estimation of inverse kernel width is based on

a subset of the original data and thus causes varying results, the change detection procedure is iterated 1,000 times for all feature sets in case of kernel k-means and spectral clustering. Concluding, the median value of accuracy in terms of Cohen's Kappa (κ) is used for final evaluation of change detection accuracy.

4. Results

After calculation and normalization of the object-based features, PCA was conducted. The number of relevant PCs was determined by means of a tripartite selection procedure. The Bartlett's test was highly significant for the PCs of all feature sets, i.e., relevant structure with respect to PCA is present in the data. The investigation of the cumulative proportion of variance resulted in two relevant PCs for each set of MS, ratio and grayscale texture features, respectively, while four PCs were needed in case of the merged features to exceed the defined threshold of 90 percent of variance. Finally, the averaged parallel analysis returned two relevant PCs for the MS features, three relevant PCs for both ratio and grayscale texture features and four relevant PCs in case of the merged feature set. According to the maximum rule of the second and third condition, two PCs were retained for the MS features, three PCs for ratio and grayscale texture features and four relevant PCs in case of the merged feature set.

Based on these relevant PCs, change detection was conducted by means of seven partitioning clustering algorithms and summarized values of resulting κ statistics are displayed in Tab. III-2. The best overall result was achieved based on the relevant PCs of the merged feature set in combination with the genetic k-means clustering algorithm. This combination shows very good classification agreement with only few errors of commission and omission, which were present at approximately equal shares (Fig. III-1). In general, the MS and merged features showed good classification results, i.e., mean κ around 0.8 for all clustering algorithms. Moderate results of mean κ around 0.6 were returned for ratio and grayscale texture features due to less favorable distributions of the classes of interest (i.e., changed and unchanged buildings) in the PC feature space.

Tab. III-2 Summarized values of kappa (κ)

Feature set	*k-means*	*PAM*	*genetic k-means*	*kernel k-means**	*spectral clustering**	*SOM*	*EM*	**Mean of clustering algorithms**
Multispectral	0.87	0.86	0.84	0.67	0.85	0.84	0.80	**0.82**
Ratio	0.73	0.73	0.66	0.41	0.57	0.67	0.06	**0.62**
Grayscale texture	0.81	0.83	0.83	0.56	0.11	0.82	0.22	**0.57**
Merged	0.90	0.89	0.91	0.78	0.82	0.90	0.78	**0.76**
Mean of feature sets	**0.83**	**0.83**	**0.81**	**0.61**	**0.59**	**0.81**	**0.47**	

* median value of 1,000 iterations

With respect to different clustering algorithms under study, superior results were achieved using k-means, PAM, genetic k-means and SOM clustering, whereas the remaining algorithms, i.e., kernel k-means, spectral clustering and the EM algorithm, returned worse results across all feature sets (Tab. III-2). Especially spectral clustering and the EM algorithm showed highly variable accuracy with respect to different feature sets with lowest accuracies of κ around 0.1. The accuracies of kernel k-means were slightly higher throughout, but showed high variation among the 1,000 iterations. In contrast, k means, PAM, genetic k-means, SOM clustering and the EM algorithm exhibited stable change detection results without any variations since they do not require estimation and initialization of crucial input parameters. As mentioned above, κ was substantially lower for ratio features across all seven clustering algorithms due to overlapping of classes in the PC feature space. Nevertheless, k-means as well as PAM achieved best results for ratio features in terms of $\kappa > 0.7$. In case of grayscale texture features, spectral clustering and the EM approach were sensitive against outliers in the PC feature space. Outliers were only present in the grayscale texture features due to few (ca. 20 objects) extraordinary strong changes from very homogeneous undeveloped surfaces at t_0-1 to highly textured buildings at t_0. The clustering algorithms PAM and genetic k means were capable of properly handling these outliers and returned good classification accuracy of $\kappa = 0.83$.

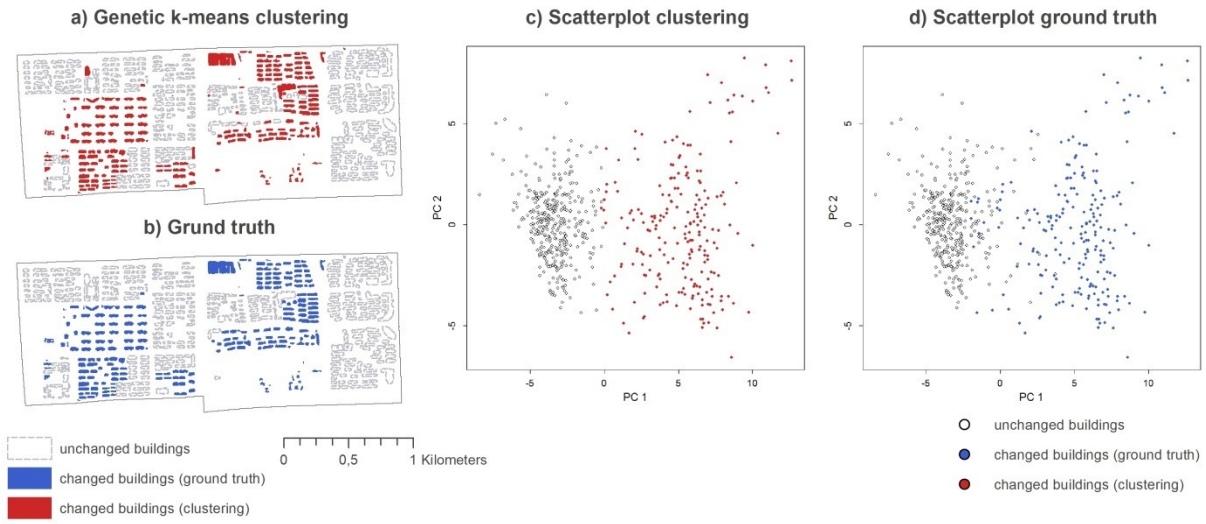

Fig. III-1 Best results of unsupervised change detection based on genetic k-means clustering using the merged feature set: a) clustering result b) ground truth map c) scatterplot of changed and unchanged buildings of PC1 against PC2 for the clustering result d) scatterplot of PC1 and PC2 of ground truth.

5. Conclusion

This study presents an unsupervised change detection approach for VHR remote sensing imagery investigating the change (i.e., growth) of the building inventory in urban areas. More specific, the aim of this study was to evaluate the capability of seven partitioning clustering algorithms under consideration of different sets of object-based input features.

The most important prerequisite for clustering algorithms employed in unsupervised change detection is that no additional input parameters besides the number of clusters are required. Although there are commonly utilized heuristics for estimation of arbitrary initialization parameters, this study showed that these unsupervised estimation procedures may cause varying classification results that are outperformed by other, more robust clustering algorithms. Furthermore, relatively simple methods like the well-established k-means algorithm provided among the best results in the presented unsupervised change detection context. Other favorable techniques were PAM, genetic k-means and the SOM clustering approach. With respect to different feature sets, the comprehensive merged feature set consisting of MS, ratio and grayscale texture features provided best results. An eligible alternative is the exclusive use of MS features, which performed similarly well in the presented unsupervised change detection analysis.

Acknowledgements

This work was supported by means of the DELIGHT project (www.delight.eoc.dlr.de, grant no. 02WCL1249A and 02WCL1249I) as well as the project EO-CITI (grant no. 033LK001A and 033LK001B), both funded by the German Federal Ministry of Education and Research, BMBF. The authors would like to thank Matthias Boes (European Space Imaging, EUSI) for providing the WorldView imagery and Dr. Klaus Martin (Company for Remote Sensing and Environmental Research, SLU) for his support.

Chapter IV. Class imbalance in unsupervised change detection – A diagnostic analysis from urban remote sensing

International Journal of Applied Earth Observation and Geoinformation 60 (2017), 83–98.

Tobias Leichtle[a, b, c], Christian Geiß[b], Tobia Lakes[c], and Hannes Taubenböck[b]

[a] Company for Remote Sensing and Environmental Research (SLU), Kohlsteiner Straße 5, 81243 München, Germany

[b] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchner Straße 20, 82234 Weßling, Germany

[c] Humboldt-Universität zu Berlin, Geography Department, Rudower Chaussee 16, 12489 Berlin, Germany

http://dx.doi.org/10.1016/j.jag.2017.04.002
Received 23rd December 2016; revised 31st March 2017; accepted 11th April 2017.

Abstract

Automatic monitoring of changes on the Earth's surface is an intrinsic capability and simultaneously a persistent methodological challenge in remote sensing, especially regarding imagery with very-high spatial resolution (VHR) and complex urban environments. In order to enable a high level of automatization, the change detection problem is solved in an unsupervised way to alleviate efforts associated with collection of properly encoded prior knowledge. In this context, this paper systematically investigates the nature and effects of class distribution and imbalance in an unsupervised binary change detection application based on VHR imagery over urban areas. For this purpose, a diagnostic framework for sensitivity analysis of a large range of possible degrees of class imbalance is presented, which is of particular importance with respect to unsupervised approaches where the content of images and thus the occurrence and the distribution of classes are generally unknown a priori. Furthermore, this framework can serve as a general technique to evaluate model transferability in any two-class classification problem. The applied change detection approach is based on object-based difference features calculated from VHR imagery and subsequent unsupervised two-class clustering using k means, genetic k-means and self-organizing map (SOM) clustering. The results from two test sites with different structural characteristics of the built environment demonstrated that classification performance is generally worse in imbalanced class distribution settings while best results were reached in balanced or close to balanced situations. Regarding suitable accuracy measures for evaluating model performance in imbalanced settings, this study revealed that the Kappa statistics show significant response to class distribution while the true skill statistic was widely insensitive to imbalanced classes. In general, the genetic k-means clustering algorithm achieved the most robust results with respect to class imbalance while the SOM clustering exhibited a distinct optimization towards a balanced distribution of classes.

1. Introduction

Change detection is one of the most intrinsic capabilities of remote sensing due to its system-inherent repetitive character of image acquisition (Singh 1989). Especially with respect to recently available satellite-based multispectral images with very-high spatial resolution (VHR), change detection still remains a challenging task and an active field of research (Bruzzone and Bovolo 2013). With increasing level of detail in VHR images,

traditional pixel-based methods for change detection become less effective and techniques from object-based image analysis (OBIA) are utilized more frequently (Hussain et al. 2013). These methods allow for characterization and associated detection of changes of entire image objects which are directly related to meaningful real-world objects. OBIA techniques are beneficial in particular over urban environments, where VHR images comprise a wealth of detail due to the large spatial heterogeneity of (mostly man-made) objects. Changes within the complex urban environment can be of various types (e.g., construction of buildings, setup of infrastructure, reconstruction of buildings, etc.), whereas the most basic use case of change detection is the dissociation of changed and unchanged areas (Ridd and Liu 1998). This two-class discrimination (i.e., binary classification of changes) is the focus in most applications of unsupervised change detection in remote sensing (Bruzzone and Bovolo 2013). Dependent on the spatial resolution of the data, object-based binary change detection applications range from classification of urban and non-urban land cover based on medium and high resolution remote sensing images (e.g., Taubenböck et al. (2012)) to discrimination of changed and unchanged buildings using VHR imagery (e.g., Tian et al. (2014)). In particular, unsupervised two-class change detection approaches focusing on buildings in VHR remote sensing images have become a vital research field (e.g., Huang et al. (2014), Wang et al. (2015)). With persistently increasing availability of VHR remote sensing imagery, automated concepts for change detection are highly required (Bruzzone and Bovolo 2013). In order to achieve high automation, unsupervised methods for change detection are preferred over supervised methods since they do not require any prior knowledge on changes between the images (Hussain et al. 2013). However, for unsupervised techniques the distribution of classes is unknown a priori and can be heavily imbalanced (e.g., the unchanged class exhibits dominantly more samples compared to the changed class in numerous applications).

From a generalized perspective, this phenomenon is referred to as the class imbalance problem (Japkowicz and Stephen 2002). The imbalanced distribution of classes deteriorates the accuracy of most standard learning and classification methods, which assume a balanced distribution of classes (Japkowicz and Stephen 2002). Learning from imbalanced data is a recent and highly active field of research in machine learning (Chawla et al. 2004; López et al. 2013). Strictly speaking, any data set that exhibits an unequal distribution of classes (i.e., in case of two classes any deviation from 50:50) may be considered as imbalanced, whereas the common definition of between-class imbalances (minority class against majority class) is in the order of 100:1 or beyond (He and Garcia 2009). Class imbalance is an intrinsic problem in many classification approaches across a large field of applications (Chawla et al. 2004). Nevertheless, from a remote sensing perspective there exist only few studies that address this phenomenon in a supervised way: A supervised neural network learning method is used in Bruzzone and Serpico (1997) for classification of agricultural land with an imbalanced distribution of classes.

Kubat et al. (1998) investigate the application of oil spill detection from radar images, which is treated as an imbalanced two-class supervised machine learning task of discriminating oil slicks (minority class) against the sea surface (majority class). The authors of Williams et al. (2009) employ a modified logistic regression approach for supervised mine classification from remotely sensed data, while García et al. (2011) present an imbalanced multi-class setting using supervised classification using hyperspectral remote sensing images. Thus, no studies were carried out based on remote sensing in an unsupervised context concerning class imbalance.

In general, there are several solutions for the class imbalance problem available in literature, which generally rely on a supervised classification strategy (López et al. 2013). According to He and Garcia (2009), a widely deployed group of solutions for imbalanced learning comprise sampling-based methods. These methods aim at balancing class proportions in the training data set prior to classification. Applicable strategies are, among others, over- and undersampling of the minority and majority classes, respectively, e.g., the synthetic minority oversampling technique (SMOTE) (Chawla et al. 2002) and its successors. Another important group of supervised solutions to the class imbalance problem are cost-sensitive learning methods. The idea behind these techniques is to develop a hypothesis that minimizes the overall costs of misclassifications in the training data set, whereas the costs of misclassifying a minority example labeled as the majority class is higher than the contrary case (Elkan 2001). Examples for these learning methods comprise cost-sensitive boosting methods (Sun et al. 2007), cost-sensitive decision trees, or cost-sensitive neural networks (Kukar and Kononenko 1998). Furthermore, kernel-based methods can be employed as supervised solutions for the imbalanced learning problem, where the transformation to a higher-dimensional feature space enables proper discrimination of the imbalanced data set (He and Garcia 2009). In addition, only relevant training samples are utilized to define the model, which may thus not be affected by imbalanced data. Examples of this technique comprise over- and undersampled support vector machines (SVMs) (Akbani et al. 2004), kernel modification methods (Wu and Chang 2005) or SVMs that include active learning strategies for efficient selection of relevant training samples (Ertekin et al. 2007). Finally, one-class learning or novelty detection methods aim at classifying only a single class (i.e., recognition-based approaches) (Pimentel et al. 2014). Examples of this approach are one-class SVMs (Lee and Cho 2006) or the autoassociator method which employs neural networks for one-class classification (Japkowicz 2001).

In contrast to these generally supervised solutions to imbalanced learning, this paper presents the class imbalance problem in an unsupervised context based on the example of urban remote sensing and change detection. In detail, the application of binary change detection of buildings using VHR remote sensing imagery is presented as an example since buildings are the most relevant and one of the most dynamic objects in complex

urban environments (Huang et al. 2014). The objectives of this paper are i) presentation of the class imbalance problem in an unsupervised change detection analysis of VHR remote sensing imagery and ii) systematic description of the nature and effects of imbalanced classes in this exemplary application setting. Thus, this study offers a sensitivity analysis framework on the nature and the effects of an arbitrary distribution of classes (i.e., class imbalance) in an unsupervised context for any two-class classification problem, which additionally enables evaluation of transferability as well as verification of the validity of any algorithm with respect to the distribution of classes.

This paper is organized as follows. Section 2 presents the two study areas, the city of Dongying in China and the city of Munich in Germany, respectively, and their individual data settings. In Section 3, a brief review of the employed methodology for change detection is given first, while Section 3.2 entails the proposed workflow for evaluation of class imbalance. Section 4 presents experimental results, highlighting the nature and the effects of imbalanced classes in Section 4.2. A detailed discussion is provided in Section 0 while concluding remarks are given in Section 6.

2. Study area and data sets

The proposed experiments on class imbalance are conducted in two disparate areas of interest with similar number of buildings, each with distinct characteristics concerning the structure of the built environment and a specific ratio of class imbalance (i.e., proportion of changed and unchanged buildings). The first experimental site is located in the dynamic Chinese city of Dongying, while the second test site is situated in the less dynamic German city of Munich. In this paper, the test site of Dongying serves as a representative of simple and less complex urban structure with regular geometric arrangement as it is present in many cities that are strongly influenced by planning. Opposed to Dongying, the morphological structure of Munich is very diverse and the buildings are arranged in a complex configuration. Such a setting is characteristic for historically developed cities. A direct comparison of structural statistics and other relevant numbers to the analyses is given in Tab. IV-1.

Tab. IV-1 Statistics of the two test sites in Dongying and Munich.

	Dongying	Munich
Number of buildings	4119	4272
Mean height of buildings [m]	14.8	14.5
Mean area of buildings [m²]	1089	497
Mean volume of buildings [m³]	17048	7247
Average distance to nearest building [m]	12.0	2.1
Main orientation of buildings	E-W (87%)	N-S (35%)
Area of test site [km²]	38.5	9.7
Built-up density [%]	17.5	29.2
Temporal scale of remote sensing images	2007 - 2013	2001 - 2010
Class proportions [unchanged : changed]	54 : 46	93 : 7

The experimental site in Dongying city comprises 4.119 individual buildings on an area of 38.5km² in 2013 (Tab. IV-1, Fig. IV-1), which were delineated from a normalized digital surface model (nDSM) with spatial resolution of 4 meters. The DSM was derived by Airbus Defence and Space based on Pléiades tristereo imagery acquired on October 18, 2013. Normalization of the DSM was performed according to Geiß et al. (2015b), while buildings were finally extracted utilizing the segmentation and classification method presented in Wurm et al. (2011). In order to achieve highest accuracy of this base data set, building geometries were checked and reworked manually. For provision of ground truth information, buildings were classified by visual interpretation according to changed and unchanged objects. Change is defined in terms of newly constructed buildings, other types of changes of buildings are absent within the experimental site. In total, 2243 buildings remain unchanged, while 1876 buildings changed within the observation period. This corresponds to a ratio of class imbalance of 54:46 (unchanged vs. changed objects), i.e., the distribution of the two classes is quite balanced in the experimental site of Dongying city. The buildings in Dongying possess an average height of about 15 meters and low built-up density (Tab. IV-1). The buildings are large (i.e., mean area of 1089 m²) with very regular, chessboard-like arrangement (Fig. IV-1). The orientation of buildings in Dongying is mainly E-W (Tab. IV-1, Fig. IV-1c) which is strongly influenced by planning. The average distance between buildings is high (Tab. IV-1) due to mostly free standing of buildings in Dongying (Tab. IV-1). Based on visual inspection, most buildings in Dongying are very similar with almost identical appearance (Fig. IV-1). Thus, the structure of the building environment can be described as relatively simple and less complex.

In case of Dongying, the remote sensing data consists of a QuickBird image from February 25, 2007 and a WorldView-2 image acquired on January 17, 2013 (Fig. IV-1), which corresponds to a temporal scale of six

years in the change detection analysis. The off-nadir view angles are 14.2° in case of QuickBird and 22.5° for the WorldView image. Both images are delivered as pan-sharpened image products by the data provider (European Space Imaging, EUSI) with four multispectral image channels (blue, green, red, and near infrared (NIR)), which were resampled to a common spatial resolution of 50cm. Both images were atmospherically corrected using the ATCOR (Atmospheric and Topographic Correction) toolbox (Richter 1996), followed by band-by-band histogram matching.

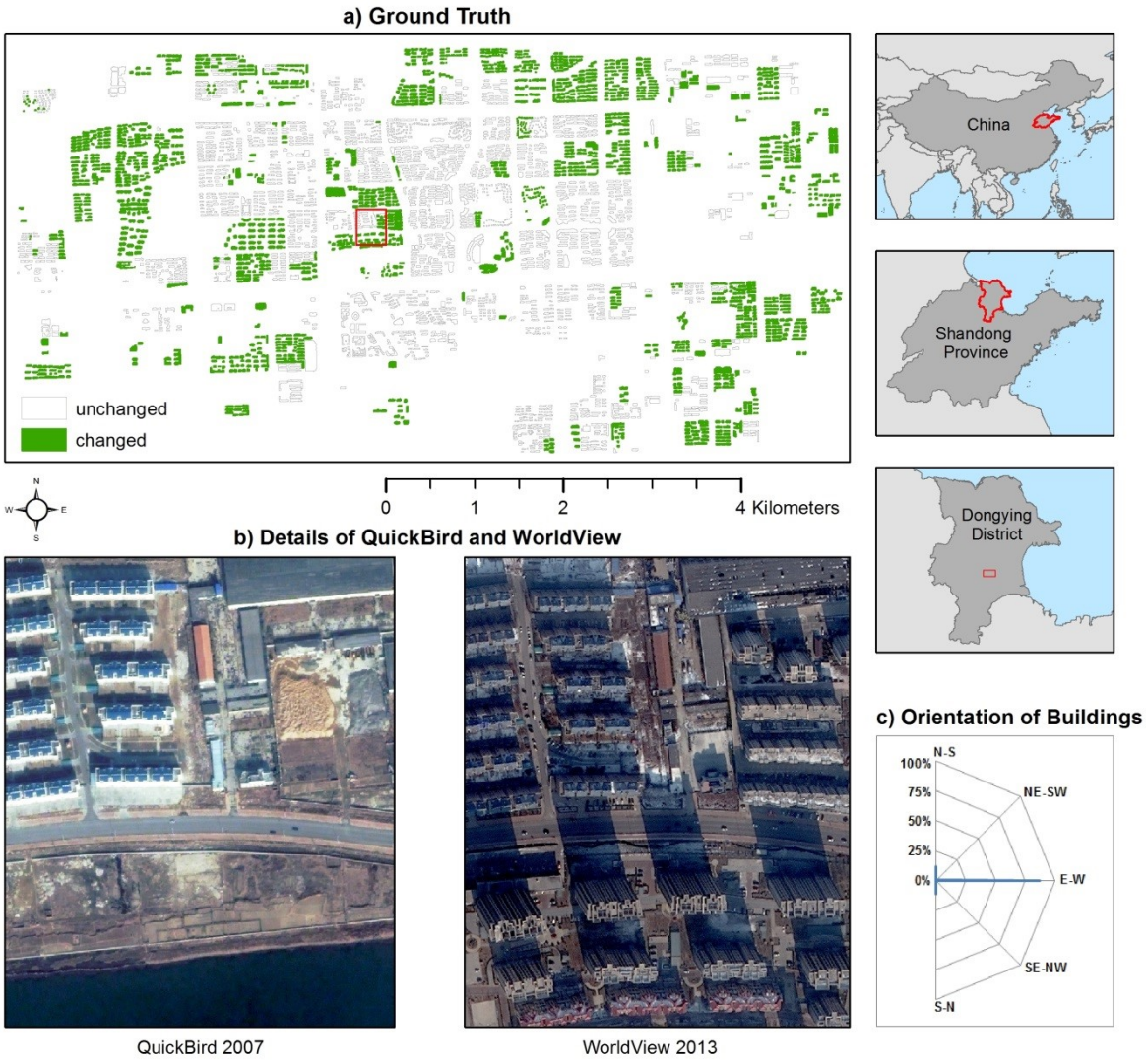

Fig. IV-1 Experimental site and available data in Dongying, China: a) Building geometries and ground truth information, b) Detailed views of multi-temporal remote sensing imagery, c) Orientation of buildings according to the main angle of polygons.

In Munich, the experimental site covers an area of 9.7km² and consists of 4272 buildings in 2010 (Tab. IV-1, Fig. IV-2). The geometries of buildings originate from volunteered geographic information, i.e., OpenStreetMap (OSM). The geometries from OSM were checked and manually corrected since the utilized remote sensing imagery is outdated (i.e., from 2010) compared to latest up-to-date OSM geometries. Ground truth information on changes was collected manually by visual comparison of the multi-temporal images. In order to provide a consistent terminology of change for both experimental sites, buildings that exhibit only slight changes were excluded from the analysis and thus, only newly constructed buildings or complete building reconstruction (i.e., demolition and new construction of buildings) was considered in case of Munich. In total, 3730 buildings remain unchanged, while 278 changed buildings were identified according to the abovementioned definition. Thus, the ratio of class imbalance for Munich is 93:7 (unchanged vs. changed objects), i.e., the distribution of the two classes is strongly imbalanced. The average height of buildings in Munich (14.5 m) is very similar to those of Dongying (14.8 m), while they are considerably larger in size (Tab. IV-1). The built-up density is higher in Munich while the distances between buildings is low (Tab. IV-1), whereas 75% of buildings are directly connected to their neighbors (distance = 0 m), which indicates perimeter development. The buildings are oriented in all directions with no significant main orientation (Tab. IV-1, Fig. IV-2c) and consist of diverse appearances (Fig. IV-2). Thus, the building alignment and general built-up structure of Munich can be described as variable and complex compared to the test site of Dongying.

The remote sensing images for change detection in Munich comprise an IKONOS image from September 3, 2001 and a WorldView image acquired on July 12, 2010, respectively (Fig. IV-2). Thus, the temporal scale for change detection in Munich is nine years. The IKONOS image possesses an off-nadir view angle of 5.5° while the off-nadir angle for the WorldView image is 13.9°. IKONOS provides pan-sharpened imagery at a spatial resolution of 1m, while WorldView resolves pixels at 50cm, whereas pan-sharpening was performed by the data provider (EUSI). Both sensors record four multispectral bands (blue, green, red, and NIR) and the spatial resolution was resampled to common pixel size of 50x50cm². Atmospheric disturbances were corrected using ATCOR (Richter 1996) and the two images were adjusted radiometrically by means of band-by-band histogram matching.

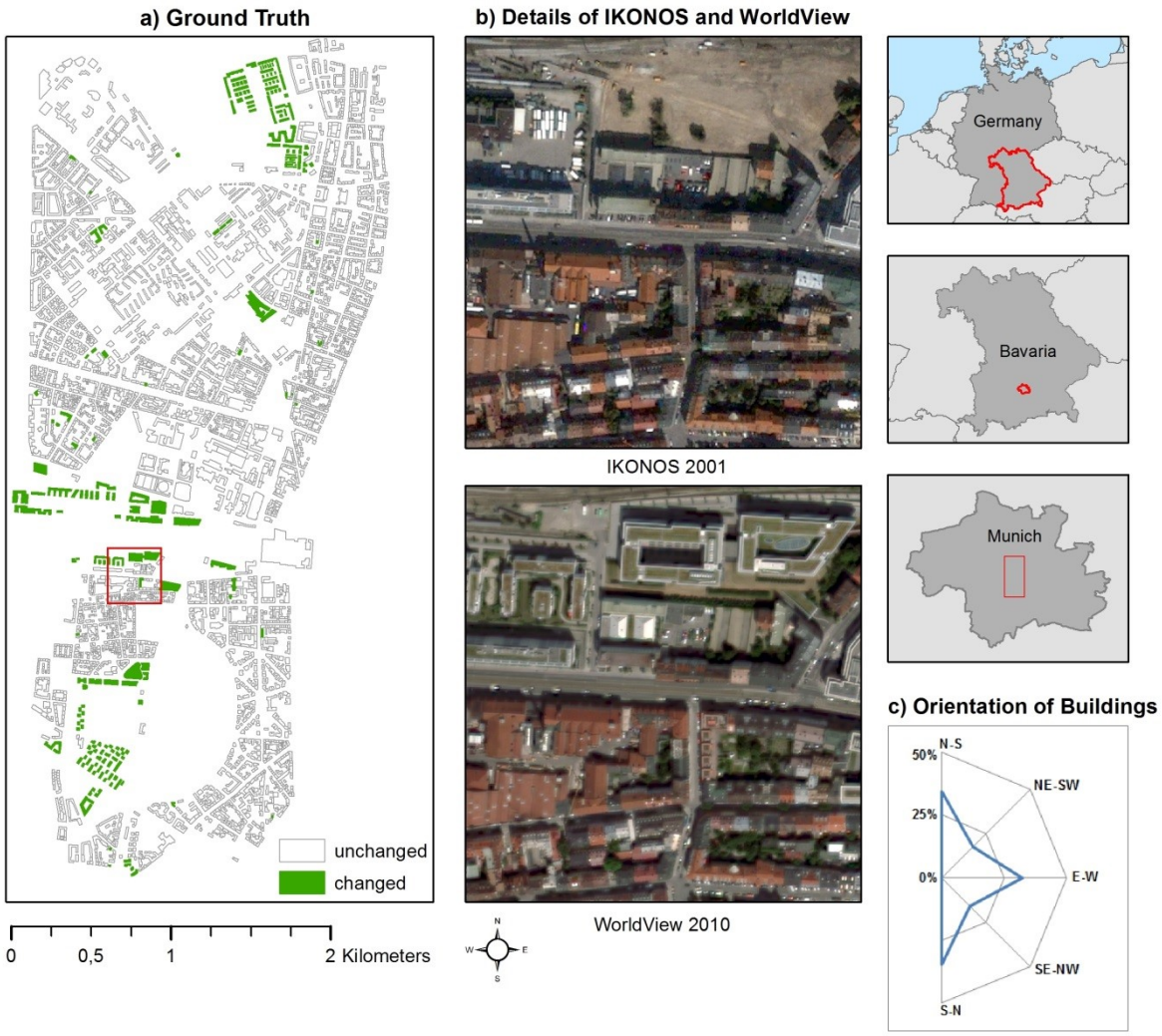

Fig. IV-2 Experimental site and available data in Munich, Germany: a) Building geometries from OSM and ground truth information, b) Details of multi-temporal IKONOS (2001) and WorldView (2010) imagery, c) Orientation of buildings according to the main angle of polygons.

Although the utilized sensors and corresponding VHR images are very similar in both test sites, there are slight differences that must be discussed. First, the spatial resolution of the original pan-sharpened images differs (WorldView 50cm, QuickBird 60cm, and IKONOS 1m) and is resampled to a common spatial resolution of 50cm in all cases. Other values of common spatial resolution might be applicable, whereas the object-based processing and the utilized features (see Section 3.1) minimize its influence. Second, also spectral properties of the sensors (i.e., range and width of spectral bands) deviate, however it can be assumed that the influence of these differences is also minimized by the utilized object-based features (see Tab. IV-2). Finally, varying viewing geometries (i.e., off-nadir angles) were found to deteriorate change detection results only slightly in case of Dongying (Leichtle et al. 2017a), whereas it can be assumed that these differences do not affect the result

significantly in Munich, since high-rise buildings (i.e., which are mostly affected by parallax distortion due to large off-nadir angles) are largely not present.

3. Methodology

The methodology of this study is based on object-based data preprocessing and unsupervised change detection (Section 3.1), which is strongly related to the unsupervised change detection approach proposed in Leichtle et al. (2017a). The framework for the diagnostic sensitivity analysis of class imbalance including a description of the deployed measures of classification accuracy is presented in Section 3.2. The complete workflow for evaluation of class imbalance is depicted in Fig. IV-3.

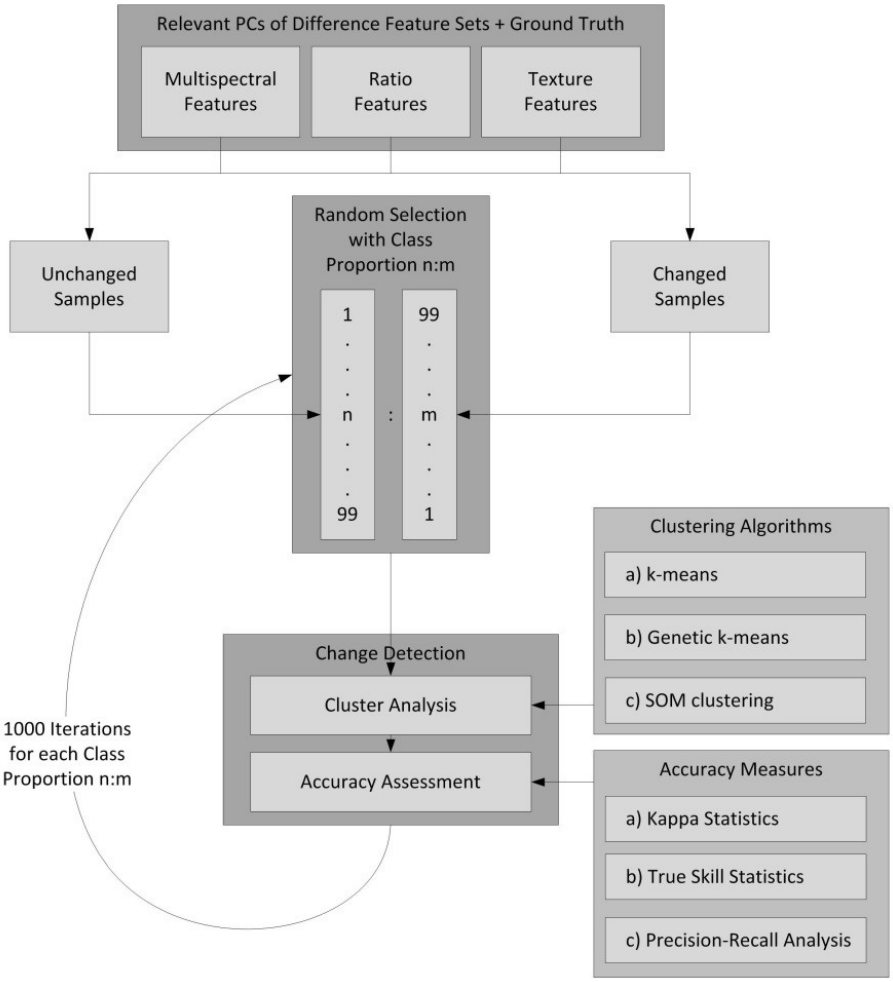

Fig. IV-3 Overview of the workflow for evaluation of class imbalance.

3.1 Preparation of object features & change detection analysis

Based on the object geometries presented in Section 2, the following object-based features are extracted from each VHR image according to Tab. IV-2. First-order statistics of the multispectral image channels, ratio features, and second-order statistics (i.e., texture features based on the grayscale image) are calculated. In addition, a merged feature set consisting of all 33 (10 multispectral, 11 ratio, and 12 texture) features is utilized in the analysis. These sets of feature have already proven their applicability for change detection of buildings in previous studies (Leichtle et al. 2017a).

Tab. IV-2 Object-based Features

Multispectral Features (10)	Ratio Features (11)	Grayscale Texture Features (12)
mean intensity of blue channel	blue / green	GLCM angular 2nd moment
mean intensity of green channel	blue / red	GLCM contrast
mean intensity of red channel	blue / NIR	GLCM correlation
mean intensity of NIR channel	green / red	GLCM dissimilarity
mean intensity of all channels (brightness)	green / NIR	GLCM entropy
standard deviation of blue channel	red / NIR	GLCM homogeneity
standard deviation of green channel	normalized green[a]	GLCM mean
standard deviation of red channel	normalized red[b]	GLCM standard deviation
standard deviation of NIR channel	normalized NIR[c]	GLDV angular 2nd moment
maximum difference between channels	NDVI[d]	GLDV contrast
	SAVI[e]	GLDV entropy
		GLDV mean

grayscale image = (blue + green + red) / 3

[a] normalized green = green / (NIR + red + green), [b] normalized red = red / (NIR + red + green), [c] normalized NIR = NIR / (NIR + red + green)

[d] NDVI = (NIR – red) / (NIR + red), [e] SAVI = [(NIR – red) / (NIR + red + L)] * (1 + L), where L = 0.5

The object-based features from each corresponding multi-temporal VHR image are combined in terms of object-based image differencing of each feature (i.e., difference feature vector) in order to implicitly include the temporal relationship for change detection. The difference features are scaled to a common range of values of [-1, 1] by means of minimum-maximum normalization.

The final processing step of feature preparation is dimensionality reduction since most of the object-based features from Tab. IV-2 are highly correlated and thus comprise redundant information. As presented in Leichtle et al. (2017a), a suitable linear feature extraction technique is principal component analysis (PCA), while for

efficient determination of the number of relevant principal a tripartite procedure can be deployed. This encompasses the Bartlett's test (Jackson 1993), a threshold condition of cumulative proportion of variance explained by the considered PCs as well as an averaged parallel analysis according to Peres-Neto et al. (2005). Provided that the Bartlett's test returns a positive result, the final number of relevant PCs is determined according to a maximum rule of the results of the cumulative proportion of variance condition and the averaged parallel analysis.

Based on the relevant PCs of each difference feature set (multispectral, ratio, grayscale texture and merged feature set), the classification of changed and unchanged buildings is conducted by means of an unsupervised clustering approach. Since the most qualified algorithm for clustering cannot be determined a priori, three suitable clustering algorithms are selected and evaluated according to Leichtle et al. (2017b-a): k-means (Jain 2010), genetic k-means (Maulik and Bandyopadhyay 2000), and the self-organizing map (SOM) clustering approach (Vesanto and Alhoniemi 2000).

For finalization of the unsupervised change detection analysis, labels are assigned to the unlabeled partitioning of the data. For the experiments on class imbalance conducted in this study, labels are assigned automatically according to the higher level of agreement according to ground truth information.

3.2 Analysis of class imbalance

The concept for evaluation of class imbalance is to generate random sets of samples with a distinct proportion of changed and unchanged objects and subsequently perform the change detection analysis iteratively (Fig. IV-3). The resulting evolution of accuracy serves as a descriptive analysis of class imbalance. The input is provided by the four sets of relevant PCs of object-based feature differences as described in Section 3.1. The input feature sets are combined with ground truth information (see Section 2) in order to separate the input data according to changed and unchanged samples.

Each iteration of change detection is based on an input data set consisting of 100 randomly selected samples. The random selection and associated change detection is repeated 1,000 times for each class proportion of n:m. n represents the proportion of samples from class 0 (i.e., unchanged buildings), whereas m is the number of randomly selected samples from class 1 (i.e., changed buildings). In order to generate a total of 100 (i.e., n + m) samples, n is increased from 1 to 99 while m is altered in a reverse manner (i.e., from 99 to 1). As mentioned in Section 3.1, the three clustering algorithms k-means, genetic k-means, and the SOM clustering approach are evaluated based on this iterative concept.

An important issue for the description and analysis of the class imbalance problem is the choice of the appropriate measure of classification accuracy (He and Garcia 2009). Typical evaluation measures are based on the confusion matrix (Tab. IV-3), which opposes predicted class labels (i.e., classification) against reference data for each class of interest.

Tab. IV-3 Layout of the confusion matrix in case of binary classification

		Reference Data	
		0 (positive)	1 (negative)
Classification	0 (positive)	TP	FP
	1 (negative)	FN	TN

TP = True Positives, FP = False Positives, FN = False Negatives, TN = True Negatives

Several measures of classification accuracy can be delineated from the confusion matrix, whereas the most straightforward descriptive metric is overall accuracy (OA) according to (1):

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

Other relevant descriptive metrics include the true positive rate (TPR according to (2), also termed sensitivity or recall), the true negative rate (TNR according to (3), also called specificity) or the positive predictive value (PPV according to (4), interchangeably named precision) (Allouche et al. 2006). The TPR and PPV are related to commonly used producers's and user's accuracies and their respective errors of commission and omission (Congalton 1991).

$$TPR = \frac{TP}{TP + FN} \text{ (sensitivity / recall)} \tag{2}$$

$$TNR = \frac{TN}{FP + TN} \text{ (specificity)} \tag{3}$$

$$PPV = \frac{TP}{TP + FP} \text{ (precision)} \tag{4}$$

Since the above mentioned singular measures of classification accuracy are often limited (Foody 2002), multivariate metrics like Kappa (κ) according to (5) and the true skill statistic (TSS) according to (6) combine both types of possible errors (i.e., omission and commission errors) and provide a more comprehensive assessment of classification performance (Allouche et al. 2006).

$$\kappa = \frac{OA - \frac{(TP+FP)(TP+FN)+(FN+TN)(TN+FP)}{(TP+FP+FN+TN)^2}}{1 - \frac{(TP+FP)(TP+FN)+(FN+TN)(TN+FP)}{(TP+FP+FN+TN)^2}} \tag{5}$$

$$TSS = TPR + TNR - 1 = sensitivity + specificity - 1 \tag{6}$$

Although κ is widely used in recent research, especially in remote sensing applications, it was proven to be sensitive to imbalanced data (Jeni et al. 2013) and thus, its utility in the presented experiments has to be evaluated carefully (Pontius and Millones 2011). To overcome this issue and in order to provide a more profound picture of class imbalance, TSS according to Allouche et al. (2006) is included as a second evaluation measure. In contrast to κ, TSS is widely insensitive to imbalanced classes (Klotz et al. 2016). Thus, both κ and TSS are evaluated in parallel across varying degrees of class imbalance.

Another technique for the analysis of classification accuracy can be realized by simultaneous consideration of two singular performance measures (i.e., (2) - (4)), which is recommended by He and Garcia (2009). A graphical solution are Precision-Recall (PR) curves, which relate precision (4) to recall (2) and are strongly related to more commonly utilized Receiver Operator Characteristics (ROC) curves. According to Davis and Goadrich (2006), PR curves are preferred over ROC curves especially in case of highly imbalanced data sets. Thus, the PR analysis is employed in this study as a third evaluation technique for class imbalance.

4. Results

Results are presented as follows. Section 4.1 shows the results of change detection for the experimental sites in Dongying and Munich, respectively. The analyses of class imbalance are conducted in Section 4.2 in order to describe the nature and the effects of class imbalance in the application context of unsupervised change detection.

4.1 Change detection results in Dongying and Munich

After preprocessing of object-based difference features (Tab. IV-2) and subsequent PCA, the number of relevant PCs was determined according to the tripartite procedure consisting of the Bartlett's test, a cumulative proportion of variance condition and an averaged parallel analysis as presented in Section 3.1. The number of relevant PCs for each experimental site is displayed in Tab. IV-4.

Tab. IV-4 Number of relevant, non-trivial Principal Components

	Dongying	Munich
Multispectral features	3	4
Ratio features	3	3
Grayscale texture features	2	4
Merged features	5	7

The number of non-trivial PCs is higher in case of Munich (Tab. IV-4) due to the greater complexity of the built-up structure and more diverse types of buildings compared to Dongying (Tab. IV-1, Fig. IV-1, Fig. IV-2). Thus, more PCs are required in order to establish an adequate representation of the input data with respect to identification of changes.

Best results of the unsupervised change detection procedure are depicted in Fig. IV-4 for the complete experimental site of Dongying and in Fig. IV-5 for the entire test site of Munich, respectively.

In case of Dongying (Fig. IV-4), superior results were achieved using the SOM clustering approach based on the grayscale texture features. In relation to the reference map, the following classification accuracy was reached:

- $\kappa = 0.87$
- TSS = 0.87
- Precision = 0.94, Recall = 0.95

These measures as well as the visual agreement of the reference and the classification maps (Fig. IV-4) show the good quality of the change detection approach in case of the balanced distribution of the target classes (i.e., unchanged and changed buildings) in Dongying. In addition, overall accuracy of 94% underlines this excellent result of change detection. The multispectral and merged feature sets as well as the clustering algorithms k-means and genetic k-means returned viable results for the test site of Dongying (κ and TSS around 0.8, precision and recall around 0.9) as well. Due to less favorable distribution in the PC input feature space (i.e., overlapping of classes), ratio features led to worse results (κ and TSS around 0.2, precision and recall around 0.6). These results can be found in detail in Tab. IV-5 in the appendix.

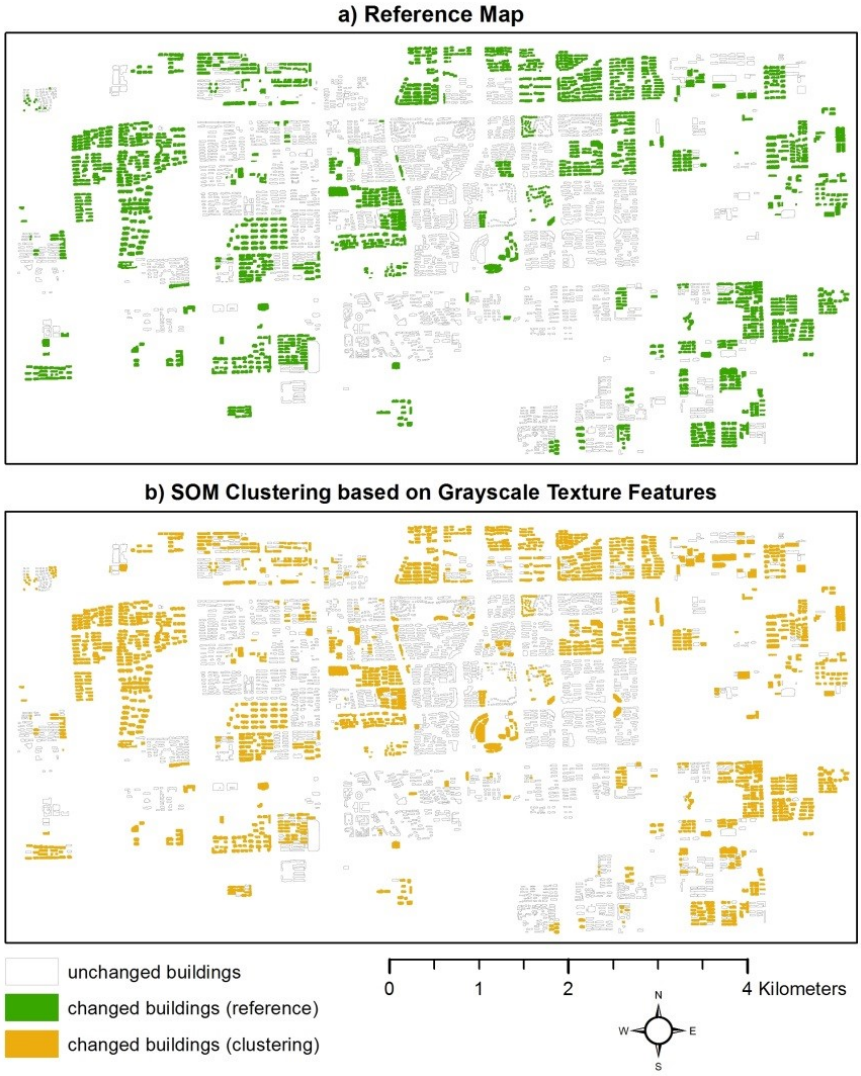

Fig. IV-4 Results of unsupervised change detection for Dongying. a) Reference map b) Best change detection result based on grayscale texture features and SOM clustering (κ = 0.87, TSS = 0.87, precision = 0.94, recall = 0.95).

For Munich, best results were reached based on k-means clustering of grayscale texture features. As visible from Fig. IV-5, the consistency of the reference map and the classification result is poor compared to the results for Dongying (Fig. IV-4), which is reflected by the calculated measures of classification accuracy:

- κ = 0.22
- TSS = 0.48
- Precision = 0.97, Recall = 0.79

Both κ and TSS show little agreement of the classification result and the reference map. The errors of classification are concretized by precision and recall, whereas the comparatively lower value of recall indicates over-classification of changed objects (i.e., high error of commission). This is due to the strongly imbalanced

distribution of classes in the experimental site of Munich (93% unchanged vs. 7% changed buildings). Note that overall accuracy of 78% for Munich is not capable of properly describing class imbalance since it does not consider the errors of over-classification of the changed class. However, k-means seeks for clusters of similar size (Jain 2010). Thus, the minority class of changed buildings is erroneously enlarged which causes a high error of commission in the classification result. This phenomenon was observed for all clustering algorithms and feature sets in the test site of Munich, whereas classification accuracy was poor (κ around 0.1, TSS around 0.3) with high error of omission (values of recall around 0.6 compared to precision around 0.9) throughout. Nevertheless, high values of precision indicate proper recognition of the changed buildings present in the test site of Munich. Complete results for Munich can be found in Tab. IV-6 in the appendix.

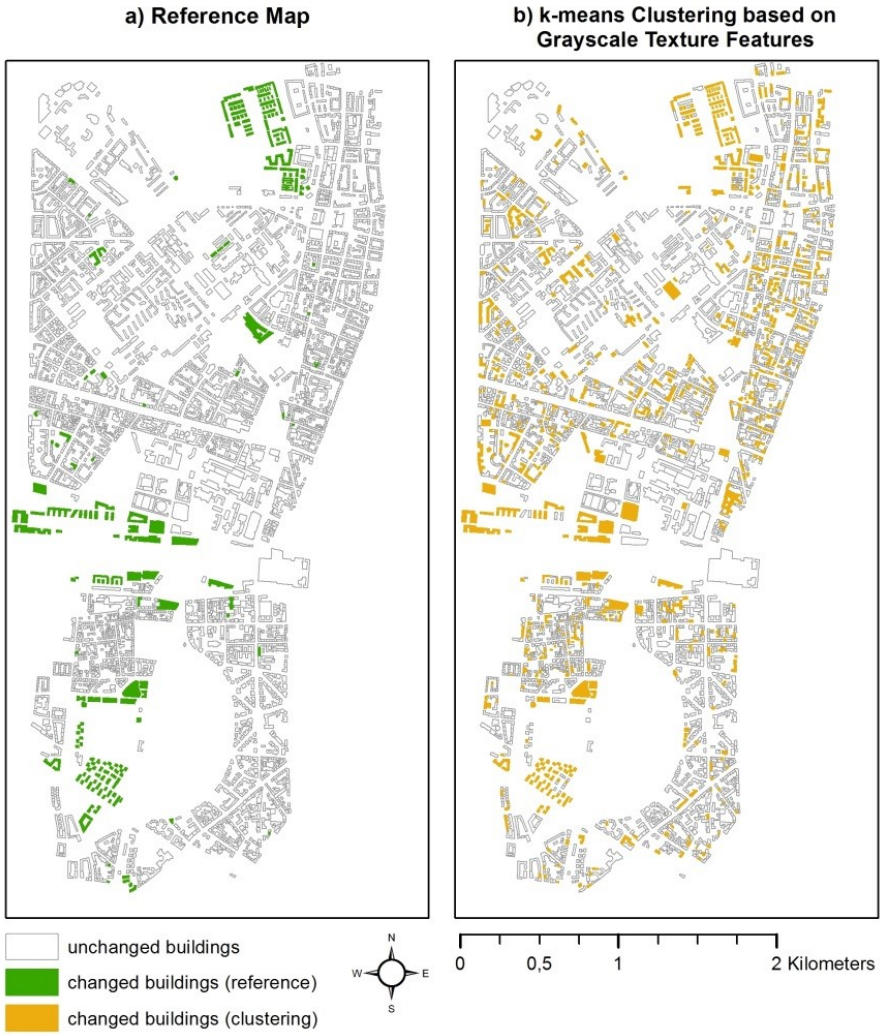

Fig. IV-5 Change detection results in case of Munich. a) Reference map b) Best change detection result based on grayscale texture features and k-means clustering (κ = 0.22, TSS = 0.48, precision = 0.97, recall = 0.79).

4.2 Description of class imbalance

Based on the workflow depicted in Fig. IV-3, different feature sets were evaluated using three clustering algorithms in order to describe the influence of imbalanced distributions of classes on classification (i.e., change detection) accuracy. In this study, classification accuracy is measured in terms of the kappa statistics (κ), the true skill statistics (TSS) and the precision-recall (PR) analysis.

Fig. IV-6 shows the results of κ statistics for the experimental site of Dongying. For each distinct class proportion of n:m (no change : change), the boxplots of Fig. IV-6 show the range of achievable classification accuracy in terms of κ within 1,000 randomly chosen data sets of sample size n + m = 100 which is spread in the order of $\kappa = \pm 0.2$ around its median value in most cases. In consistency with the original change detection results presented in Section 4.1 (red dotted line in Fig. IV-6), ratio features performed generally worse across the complete range of class distributions, i.e., highest values of κ were only around 0.5. Accordingly, all other feature sets reach very good accuracies in terms of κ well above 0.6 for all three clustering algorithms. Nevertheless, relatively highest accuracies can be observed throughout at more or less balanced distributions of classes as in case of the real-world example of the dynamic city of Dongying. In contrast, accuracies converge towards $\kappa = 0.0$ for extremely imbalanced data settings, which occur on both sides of the range of degrees of class imbalance, i.e., at n:m = 1:99 and n:m = 99:1, respectively. Additional conclusions can be drawn from the kurtosis of the resulting plots which is of rather pointed shape (i.e., leptokurtic) in case of SOM clustering opposed to the form of a plateau (i.e., platykurtic) for genetic k-means clustering. In this context, kurtosis describes the degree of optimization of clustering algorithms towards a balanced distribution of classes. SOM clustering returns best results within a narrow range of degrees of class imbalance while genetic k-means is suitable for a broad range of class imbalance from approximately n:m = 20:80 to n:m = 80:20 for all feature sets.

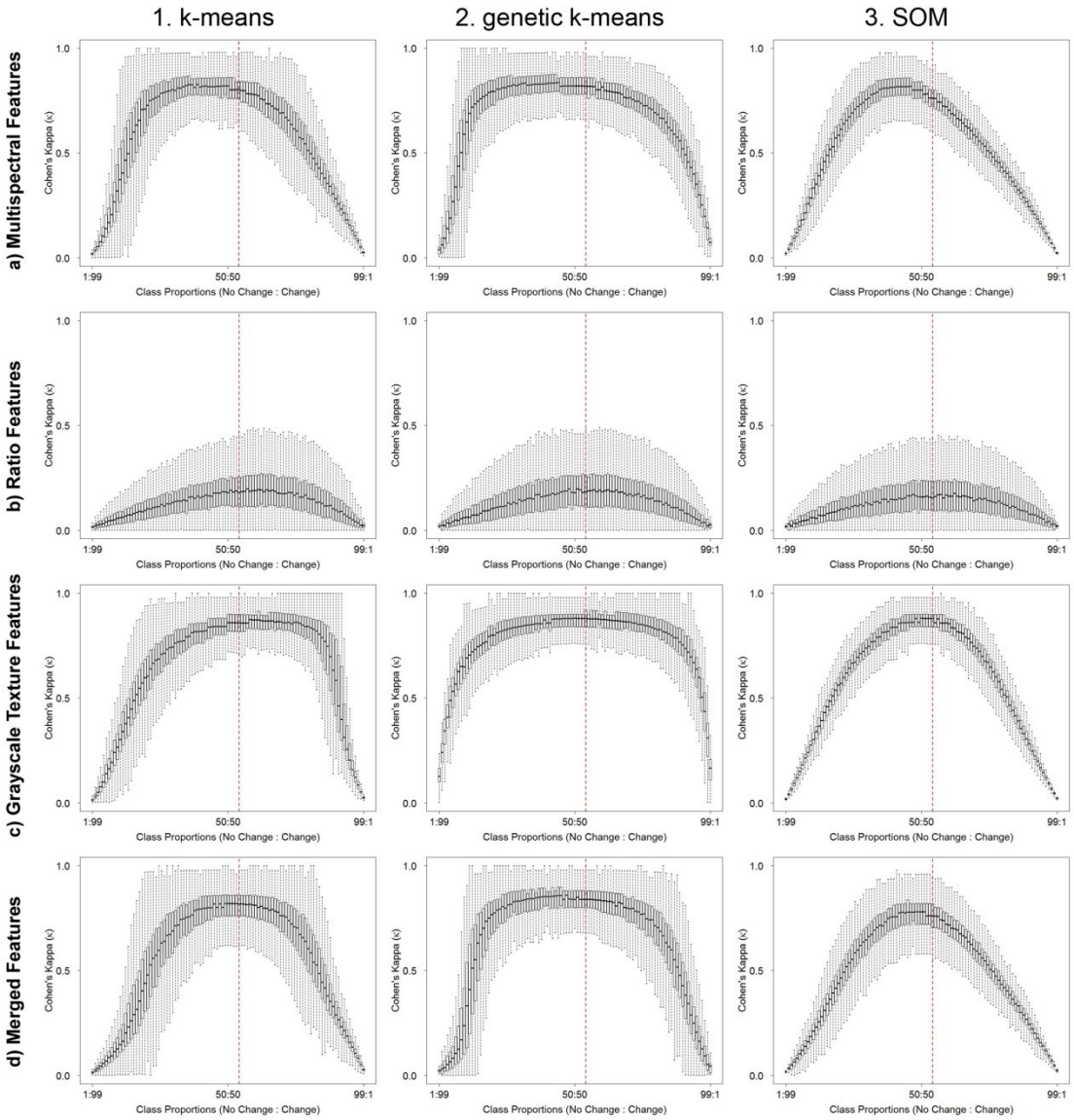

Fig. IV-6 Boxplots of κ across possible degrees of class imbalance (n:m, 1,000 iterations each) in case of Dongying: a) Multispectral features b) Ratio features c) Grayscale texture features d) Merged features. Column 1 shows results for k-means, column 2 depicts change detection accuracy according to genetic k-means, column 3 displays the performance after SOM clustering. The red dotted line shows class imbalance of 54:46 corresponding to the original classification (Fig. IV-4).

In contrast to the results of Dongying, values of κ are generally lower in case of Munich (Fig. IV-7), which is mainly induced by the greater complexity of the building structure and more diverse types of buildings in this test site (see Section 2). Nevertheless, this cannot be observed throughout: ratio features perform superior in the experimental site of Munich compared to Dongying which can be explained by the diverse types of buildings present in the test site of Munich that are represented more adequately by means of ratio features. Another

difference in comparison to Dongying is the greater range of achievable accuracy within 1,000 randomly chosen data sets in the order of $\kappa = \pm 0.3$ around its median value. Again, this can be justified by the greater complexity and diversity of the built-up structure in Munich and thus, correct identification of changes becomes more challenging. As mentioned in Section 2, the distribution of classes is strongly imbalanced (n:m = 93:7) for the original classification of Munich and hence, classification accuracies are poor for the original data set (Tab. IV-6a) and its specific distribution of classes (red dotted line in Fig. IV-7). In case of relatively balanced class distributions, the unsupervised detection of changes reaches good values of κ in the order of 0.6 depending on the utilized feature set and clustering algorithm. For Munich, k-means returned best results for most feature sets. Although not visible as clearly compared to the results of Dongying, the kurtosis of the SOM clustering shows optimization towards a balanced distribution of classes (i.e., kurtosis is leptokurtic) while genetic k-means returns good results across a broader range of class distributions (i.e., platykurtic kurtosis). With respect to the shape of the plots, especially k-means and the SOM clustering possess noticeable skewness in terms of higher absolute κ towards greater proportion of changed building samples. Simultaneously, the range of values of κ widens. This effect can be explained by the small absolute number of changed samples (278) in the test site of Munich which show very diverse characteristics with respect to the utilized feature sets. Thus, the classification accuracy of changed buildings is strongly dependent on the (random) choice of samples from its rather small population.

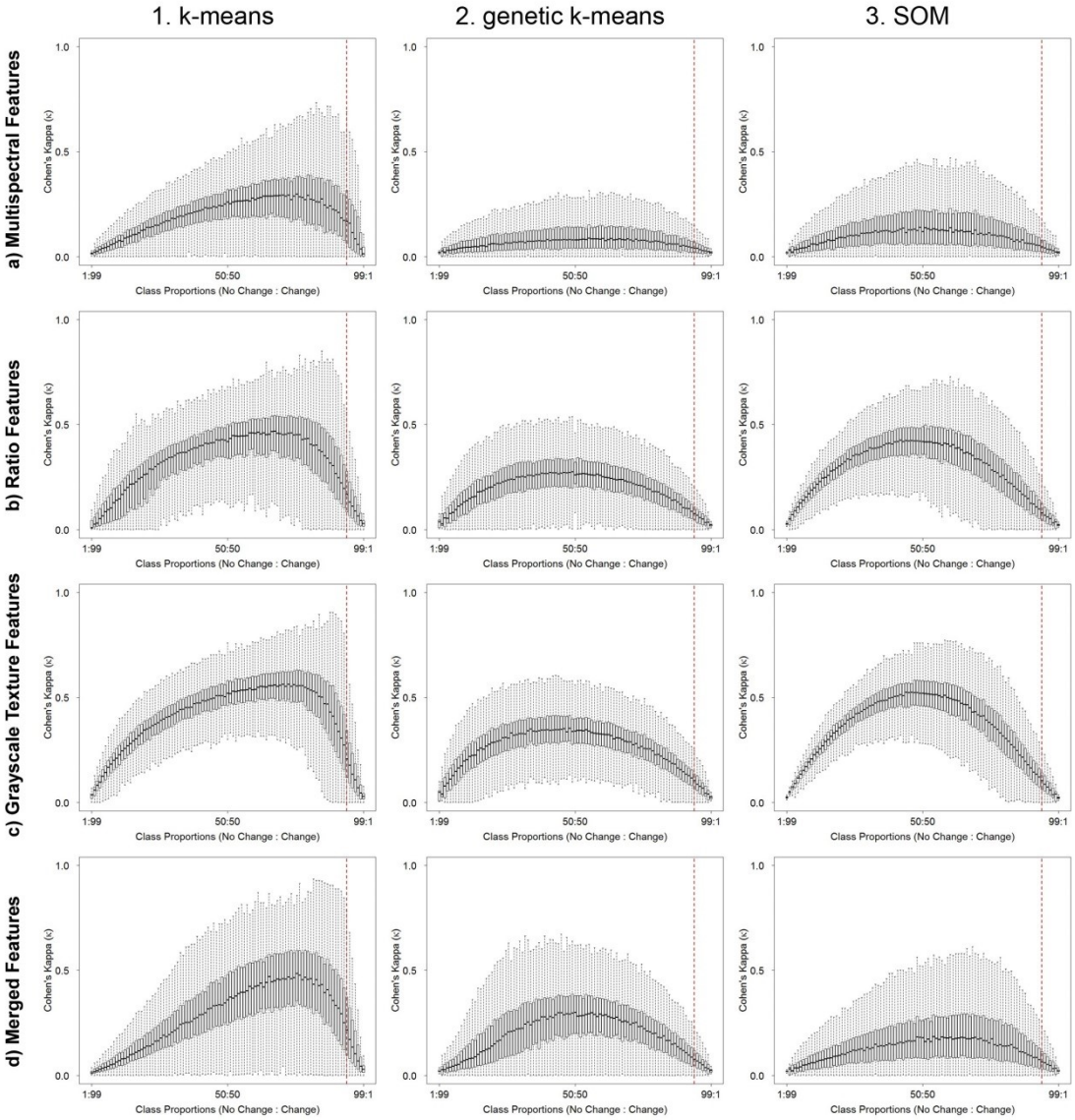

Fig. IV-7 Accuracy of change detection according to κ across possible degrees of class imbalance (n:m, 1,000 iterations each) in Munich: a) Multispectral features b) Ratio features c) Grayscale texture features d) Merged features. Column 1 shows results for k-means, column 2 depicts change detection accuracy according to genetic k-means, column 3 displays the performance after SOM clustering. The red dotted line shows class imbalance of 93:7 corresponding to the original classification (Fig. IV-5).

Fig. IV-8 depicts the mean values of TSS for both test sites and all combinations of feature sets and clustering algorithms. Notably, TSS possesses values of around 0.5 for strongly imbalanced data settings (i.e., at n:m = 1:99 and n:m = 99:1, respectively) in most cases. This behavior of TSS is inherent, since TSS is equally dependent on the classification accuracy for the minority and the majority class, i.e., TPR and TNR, respectively (see Eq. 6), whereas for strongly imbalanced settings one of them tends to 1 while the other tends to 0. Thus, the tendency of all graphs in Fig. IV-8 towards a minimum or a maximum in the direction of more balanced class distributions is determined according to the inherent threshold value of TSS = 0.5. An exception to this effect occurs in case of Dongying using the genetic k-means algorithms together with texture features whereas this combination returns excellent results with respect to TSS well above 0.8, also for strongly imbalanced data settings. This data setting is the best representation of unchanged and changed buildings which is properly identified by genetic k-means clustering. In general, TSS emphasizes the skewness across different class distributions due to individual classification performance of the two classes and their proportions, since TSS is widely insensitive to imbalanced classes. In Fig. IV-8 this is revealed by means of asymmetric shape of TSS across the range of class distributions. The asymmetry of TSS is also related to the shape of median values of κ in Fig. IV-6 and Fig. IV-7, respectively. In this regard, genetic k-means is generally less sensitive towards the distribution of classes, while k-means and SOM clustering tend to greater influence (Fig. IV-8). This behavior can be accounted for by the individual classification performance of unchanged and changed buildings as in case of κ statistics.

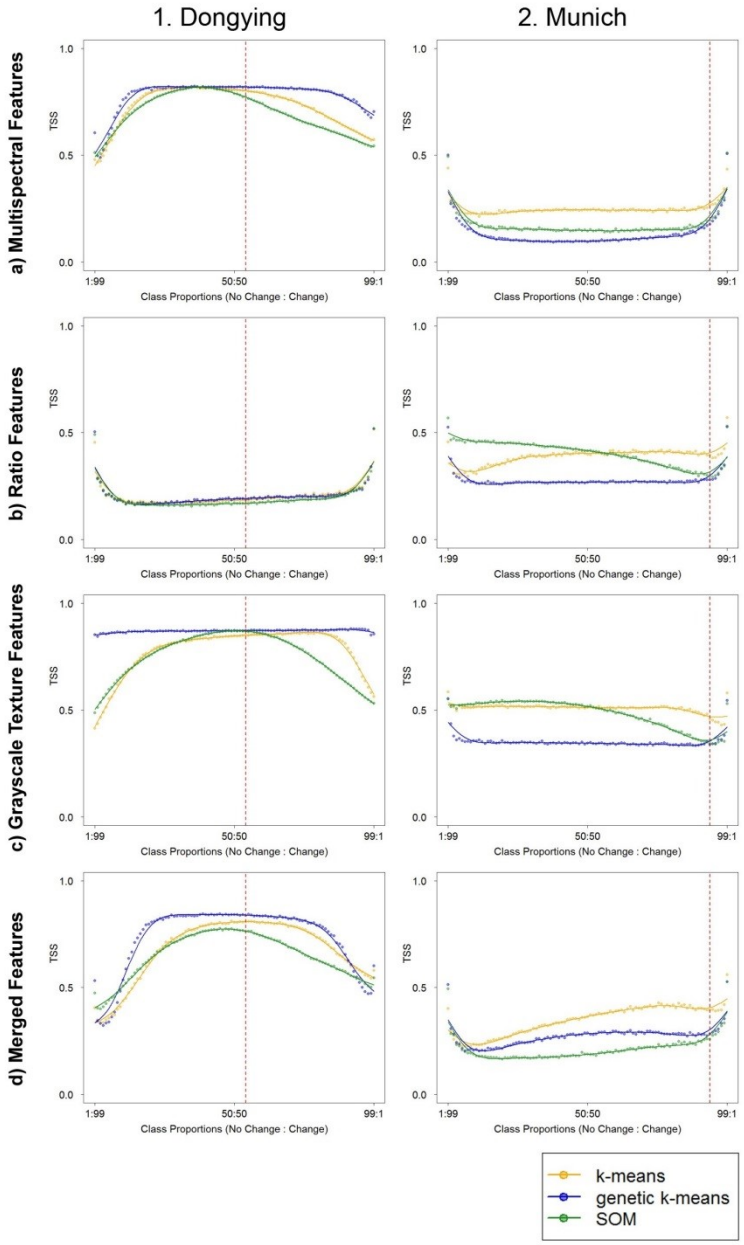

Fig. IV-8 True Skill Statistics (TSS) across possible degrees of class imbalance (n:m, mean value of 1,000 iterations each) for test sites in Dongying (left) and Munich (right): a) Multispectral features b) Ratio features c) Grayscale texture features d) Merged features. Colors refer to different clustering algorithms (k-means, genetic k-means and SOM). The red dotted line indicates the distribution of classes for the two test sites.

In the PR analysis, the range of class imbalance is reproduced in terms of precision around 0 and recall approximating 1 at n:m = 1:99, whereas towards n:m = 99:1 precision evolves towards 1, while recall approximates 0.5 (Fig. IV-9). The tendency of recall towards 0.5 can be explained by the high number of false negatives for the strongly imbalanced data setting at n:m = 99:1 (see Eq. 2). In the PR analysis, the sensitivity to imbalanced classes appears by means of slow increase of precision, whereas robust combinations of clustering algorithms and feature sets possess a rapid ascent of precision (Fig. IV-9). In general, best results of PR analysis are indicated by approximation of high values of precision and recall, i.e., allocation of the graph in the top right corner of the PR space. This is realized by all feature sets except ratio features in the test site of Dongying, while in case of Munich, relatively best results are achieved by k-means clustering across all feature sets (Fig. IV-9). With respect to imbalanced classes (i.e., evolution of the PR-curves from point A to B in Fig. IV-9), discontinuous decrease of recall indicates superior classification accuracy of one of the two classes. This especially applies to k-means clustering in the test site of Munich, where changed buildings were captured less accurate due to the greater complexity of the building structure and more diverse types of buildings. This effect increases with rising proportion of samples of this class (i.e., changed buildings), which was also indicated by skewness of κ (Fig. IV-7) as well as in behavior of TSS (Fig. IV-8). The rate of increase towards highest values of precision indicates optimization towards a balanced distribution of classes, whereas this rate is especially high for genetic k-means clustering for several feature sets in both test sites (Fig. IV-9). This behavior is related to good classification performance across a broad range of class distributions, which is also consistent with the kurtosis of κ in Fig. IV-6 and Fig. IV-7.

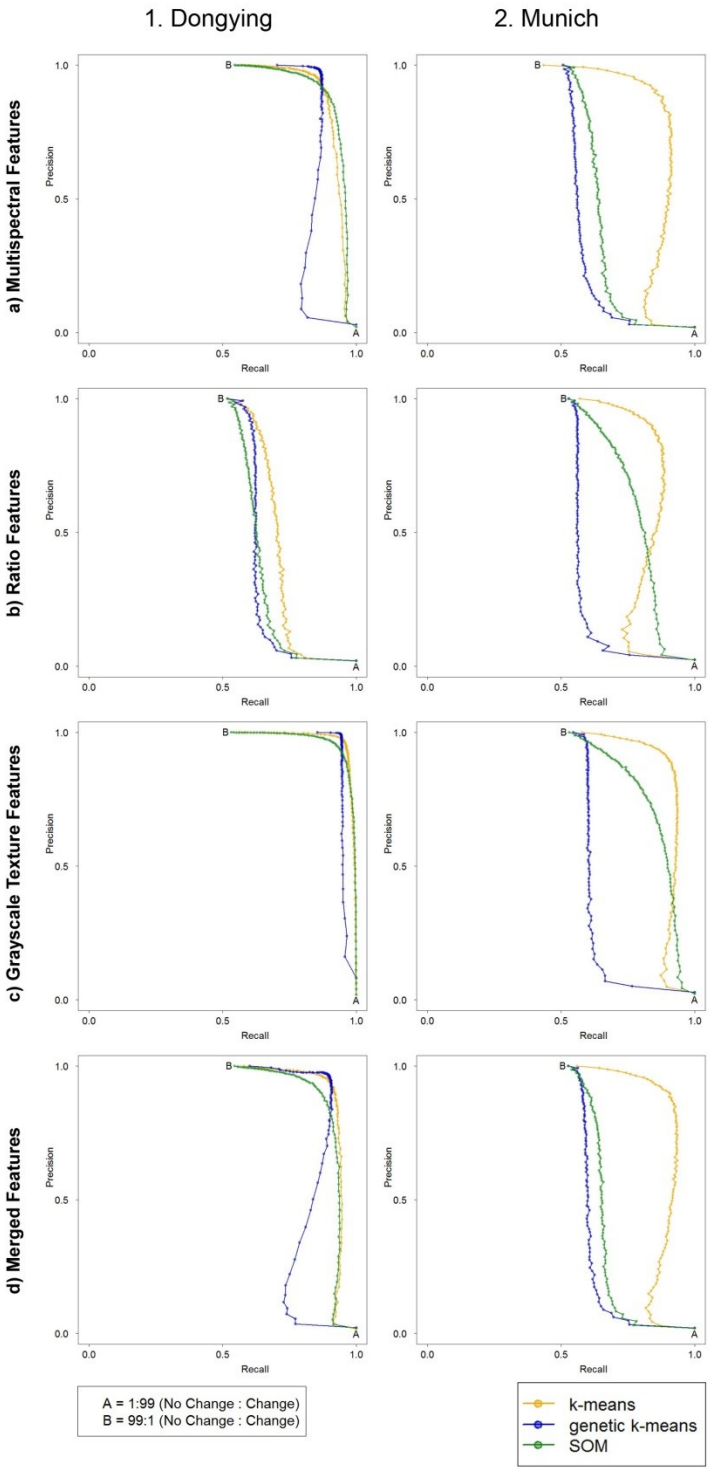

Fig. IV-9 Precision-Recall (PR) curves of unsupervised change detection across possible degrees of class imbalance (n:m, mean value of 1,000 iterations each) in Dongying (left) and Munich (right): a) Multispectral features b) Ratio features c) Grayscale texture features d) Merged features. Colors refer to different clustering algorithms (k-means, genetic k-means and SOM).

5. Discussion

This paper presents a characterization of class imbalance with respect to the exemplary remote sensing application of unsupervised change detection in urban areas. Previous analyses on remote sensing related to class imbalance (Bruzzone and Serpico 1997; García et al. 2011; Kubat et al. 1998; Williams et al. 2009) present only single case studies, often aiming at supervised methodological solutions without a detailed and systematic assessment of the nature and the effects of imbalanced classes. The supervised learning technique for multiple imbalanced classes proposed in Bruzzone and Serpico (1997) achieves classification errors in the order of 1 - 21% dependent on the considered classes and their distribution within the feature space. The imbalanced two-class classification application of oil spill detection in Kubat et al. (1998) shows accuracies of around 70% while the authors conclude that further research in the field of classification of imbalanced data is urgently required. Williams et al. (2009) reach classification performance of up to 0.9 using logistic regression but state that this method is only effective if the minority class is well represented within the feature space. Finally, García et al. (2011) employ SMOTE and random undersampling for handling a set of 16 imbalanced agricultural classes in hyperspectral images reaching accuracies of around 0.7.

In the current study, very good accuracies in terms of $\kappa = 0.87$ (TSS = 0.87, precision = 0.94, recall = 0.95) were achieved in the test site of Dongying with balanced distribution of classes, while in the imbalanced data setting of change detection in Munich, κ was 0.22 (TSS = 0.48, precision = 0.97, recall = 0.79). In contrast to previous research with respect to remote sensing, this paper focusses on the description and systematic sensitivity assessment of imbalanced classes in an unsupervised change detection application. The proposed diagnostic framework further enables evaluation of transferability as well as verification of the validity in any two-class classification problem. This study revealed that the absolute classification accuracy that can be achieved is mostly dependent on the input feature set while the employed clustering algorithm is only a secondary factor in this regard. Most robust results were achieved by a merged feature set which incorporated multispectral, ratio, and texture features. A reasonable alternative is the exclusive use of texture features based on the grayscale image which performed almost equivalently well. With respect to imbalanced class distributions, the utilized clustering algorithms possessed highest influence on change detection accuracy. Results confirmed the sensitivity of k-means against outliers and unfavorable data distributions (Xu and Wunsch 2010), which resulted in decreasing accuracy towards imbalanced distributions of classes. Genetic k-means led to superior results across a wide range of degrees of class distributions due to its inherent capability of detecting

clusters of arbitrary shape as well as its robustness in general (Xu and Wunsch 2010). Least robust results were achieved by the SOM clustering approach, which tended to show best results only in case of balanced class distributions which can be justified by its sensitivity to outliers according to Xu and Wunsch (2010). In addition, different measures of classification accuracy were found to respond differently to imbalanced classes. The most significant response is shown by κ, while TSS was widely insensitive to the conducted experiments of class imbalance. Throughout all analyses, genetic k-means demonstrated highest robustness with respect to imbalanced data settings, i.e., good classification performance was reached in the order of approximately n:m = 20:80 to n:m = 80:20 for all feature sets. Although the original k-means algorithm seeks for clusters of similar size (Jain 2010), it was also found suitable for a wide range of degrees of class imbalance. In contrast, the SOM clustering showed a distinct tendency of optimization towards a balanced distribution of classes. Thus, the choice of clustering algorithm is crucial when dealing with unsupervised methods in applications where imbalanced class distributions can possibly occur.

Nevertheless, from a more generalized perspective it must be considered that the proportions of changed and unchanged objects (e.g., buildings in urban areas) in remote sensing imagery is dependent on the type of application, geographical setting and temporal scale of the basic two-class change detection problem in urban areas (Lu et al. 2014a). *First*, the type of application determines the distribution of classes in change detection of buildings. In this context, the field of damage detection after natural hazards (e.g., earthquakes (e.g., Dell'Acqua and Gamba (2012), Taubenböck et al. (2013)), (tropical) cyclones (Radhika et al. 2015), tsunamis (e.g., Bovolo and Bruzzone (2007a), Olsen et al. (2013)), etc.) can be considered opposed to monitoring the process of urban growth (e.g., Huang et al. (2014), Tian et al. (2014), Wang et al. (2015), etc.) with respect to the distribution of changed and unchanged buildings. In case of urban growth, buildings are constructed continuously within a specific period of time, while the effect of natural hazards is revealed immediately in remote sensing imagery. In the latter case, longer time frames are likely to deteriorate the change detection result due to building reconstruction. *Second*, the ratio of changed and unchanged buildings is additionally influenced by the geographic location of a city, which is directly related to the degree of development of its home country. For example, natural hazards tend to be much more destructive in developing countries due to high vulnerability in these areas (Birkmann et al. 2013). On the other hand, cities in developing and emerging countries are mostly much more dynamic compared to urban areas in the industrialized Western world with respect to urban growth (Taubenböck et al. 2012). In addition to these two determinants, the *third* factor influencing the amount of changing buildings is the temporal scale of the remotely sensed image pairs. Naturally, the magnitude of urban growth increases with time, while most applications of damage detection after natural hazards are directly related

to a specific moment in time. In the latter application, longer temporal scales complicate accurate detection of changes (i.e., hazard-induced damage) since non-related processes (e.g., reconstruction of buildings subsequent to the hazard, regular construction of buildings, etc.) become visible in the multi-temporal imagery. These three determinants control the proportions of changed and unchanged buildings in urban change detection settings. These factors must be considered especially with respect to unsupervised remote sensing techniques, as presented in the exemplary application of change detection of buildings.

In this study, two exemplary test sites of different characteristics were analyzed. In both cases, the application context is urban growth while temporal scales are comparable (six years in case of Dongying and nine years for Munich). Obviously, the construction activity in the strongly emerging economy of China is much more dynamic compared to the industrial nation of Germany. Therefore, despite the slightly longer temporal scale for the test site in Munich, Dongying exhibits a much higher proportion of changed (i.e., newly constructed) buildings and thus, a balanced distribution of classes. From a perspective of even longer temporal scales (i.e., several decades), the distribution of classes would shift to a balanced situation in Munich while Dongying would possess an imbalanced distribution with a majority of changed buildings.

6. Conclusion

This study presents the class imbalance problem in the exemplary application of unsupervised change detection of buildings in VHR remote sensing imagery. Two test sites of very different characteristics, i.e., with balanced and imbalanced distribution of classes, respectively, as well as distinct built-up structures were analyzed exemplarily to illustrate the influence of imbalanced classes. In this application context of change detection in urban areas, considerable effects on classification performance were confirmed dependent on the distribution of classes. Furthermore, a descriptive and systematic diagnostic sensitivity analysis of the nature and effects of an imbalanced distribution of classes was conducted. The systematic evaluation of a large range of possible degrees of class imbalance revealed that the employed object-based feature set controls the absolute accuracy of change detection while the utilized clustering algorithm possesses the highest influence on classification performance in both test sites dependent on the distribution of classes. Concerning different measures of accuracy, kappa showed the most significant response to class distribution while TSS was widely insensitive to imbalanced classes. With respect to the three clustering algorithms under investigation, genetic k-means returned best and most robust classification results for degrees of class imbalance of 20:80 to 80:20. Also

the original k-means algorithm revealed good suitability across a wide range of class distributions while the SOM clustering exhibited a distinct optimization towards a balanced distribution of classes. Thus, dependent on the utilized measure of classification accuracy, the choice of a proper representation of the data (i.e., feature set) as well as the application of a suitable clustering algorithm must be considered carefully especially in an unsupervised application context as presented in this study.

In this regard, the proposed methodology can serve as a diagnostic analysis framework to evaluate transferability with respect to an arbitrary distribution of classes in any two-class classification problem. In this context, also other classification and change detection approaches can be evaluated based on the proposed framework for sensitivity analysis of class imbalance. Besides the analysis of binary changes at object level, also other relevant properties (e.g., object size, refined types of change) may be investigated and possibly included in order to enhance implications of imbalanced classes. Notably, an extension of the methodology to three or more classes must be established and investigated in future research. Besides supervised methods available in literature, unsupervised concepts are needed in order to improve classification accuracy in case of strongly imbalanced data settings. This methodological extension could enlarge possible applications of unsupervised change detection to more imbalanced data settings, e.g., analysis of urban growth in less dynamic cities over relatively short time scales like the case study of Munich presented in this paper.

Appendix

Tab. IV-5 Accuracy measures for change detection in the experimental site of Dongying. a) Kappa b) True Skill Statistics.

a) Kappa

	k-means	genetic k-means	SOM
Multispectral features	0.81	0.81	0.79
Ratio features	0.18	0.18	0.18
Grayscale texture features	0.86	0.87	0.87
Merged features	0.82	0.84	0.80

b) TSS

	k-means	genetic k-means	SOM
Multispectral features	0.81	0.82	0.80
Ratio features	0.18	0.18	0.18
Grayscale texture features	0.85	0.87	0.87
Merged features	0.82	0.85	0.81

Tab. IV-6 Accuracy measures for change detection in the experimental site of Munich. a) Kappa b) True Skill Statistics.

a) Kappa

	k-means	genetic k-means	SOM
Multispectral features	0.15	0.01	0.01
Ratio features	0.14	0.08	0.07
Grayscale texture features	0.22	0.09	0.09
Merged features	0.11	0.07	0.06

b) TSS

	k-means	genetic k-means	SOM
Multispectral features	0.27	0.06	0.05
Ratio features	0.37	0.26	0.25
Grayscale texture features	0.48	0.33	0.33
Merged features	0.29	0.24	0.20

Acknowledgements

This research was conducted in the context of the DELIGHT project (www.delight.eoc.dlr.de, grant no. 02WCL1249A and 02WCL1249I) as well as the project EO-CITI (grant no. 033LK001A and 033LK001B), both funded by the German Federal Ministry of Education and Research, BMBF. The authors would like to thank Matthias Boes (European Space Imaging, EUSI) for providing the WorldView imagery and Dr. Klaus Martin (Company for Remote Sensing and Environmental Research, SLU) for his support. The authors would also like to thank the anonymous reviewers for their very helpful comments and suggestions.

Chapter V. Has Dongying developed to a ghost city? - Evidence from multi-temporal population estimation based on VHR remote sensing and census counts

Computers, Environment and Urban Systems 78 (2019), 101372.

Tobias Leichtle[a, b], Tobia Lakes[b], Xiao Xiang Zhu[c,d], and Hannes Taubenböck[a]

[a] German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Münchner Straße 20, 82234 Weßling, Germany

[b] Humboldt-Universität zu Berlin, Geography Department, Rudower Chaussee 16, 12489 Berlin, Germany

[c] German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Münchner Straße 20, 82234 Weßling, Germany

[d] Technical University of Munich (TUM), Signal Processing in Earth Observation (SiPEO), Arcisstraße 21, 80333 Munich, Germany

https://doi.org/10.1016/j.compenvurbsys.2019.101372
Received 05th February 2019; revised 16th June 2019; accepted 21th July 2019.

Abstract

With ongoing growth and continuous development of cities, the world is turning into an urban society. In this context, urbanization, population growth and migration towards urban areas are global trends. These processes are highly dynamic especially in China, with highest rates of urbanization worldwide. In contrast to these well-known trends, a recently emerging and rarely studied side effect is “ghost cities”, which became a notable phenomenon in China recently. A ghost city is commonly defined as a new urban development that is running at severe undercapacity with respect to population and businesses and where the availability of housing and public infrastructure significantly exceeds the practical demand. Against this background, this study presents a framework based on remote sensing for the assessment of the presence or absence of the ghost city phenomenon in a typical highly dynamic Chinese city. For this purpose, remote sensing data with very-high resolution (VHR) are employed for establishment of a 4d functional city model and subsequent population capacity estimation based on a statistical approach. Regarding the components of the functional 4d city model, the multi-temporal building model and classification of building types associated with residential and non-residential function returned very high accuracies with κ of 0.73 and 0.89, respectively, while the number of floors was estimated with coefficient of determination of 0.91. Compared to the numbers of official census counts, the multi-temporal population capacity estimation revealed a considerable mismatch of available living space based on VHR remote sensing data and actual population counts. According to the conceptual framework of this study, this disagreement indicates a high likelihood and significant evidence for the emergence and presence of the ghost city phenomenon for the city of Dongying. In addition, a detailed spatial assessment was conducted in terms of an index comparing the dynamics of residential developments and population numbers to provide an impression of specific regions of the urban area which are most likely to suffer from the ghost city phenomenon.

1. Introduction

Urbanization, population growth and displacement of population concentration towards cities are global trends. These processes keep turning the world into an urban society as the largest parts of the additional population is projected to become urban dwellers (United Nations 2014). This trend is especially dynamic in China, with highest urbanization rates worldwide (Seto et al. 2011; Taubenböck et al. 2014). The urban share of

the total population increased by almost 30 percent within the past three decades, and growth rates are projected to persist in the future (World Bank and Development Research Center of the State Council 2014). With respect to the physical expansion of urban areas, the spatial extent of urban land in China has tripled during the past 30 years while the expansion reached a factor of five in the dynamic coastal provinces in the South-East of China (Schneider and Mertes 2014).

In this regard, remote sensing has proven an adequate tool for quantitative measurement and continuous observation of the urban built-up extent and its spatial expansion at global (Schneider et al. 2010; Taubenböck et al. 2012) as well as at national scale, e.g., with respect to China (Li and Gong 2016; Shi et al. 2017; Xie and Weng 2016). In literature, numerous methods for mapping and monitoring of urban built-up extent exist using various types of remote sensing data at different spatial scales. For example, Schneider et al. (2010) utilize 500m MODIS data for global classification of urban land cover, while the authors of Taubenböck et al. (2012) employ multi-temporal Landsat and TerraSAR-X imagery for monitoring of megacities worldwide at a spatial resolution of 30m. Another data source is multi-temporal nighttime light imagery, which is used for global mapping of urbanization dynamics (Ma et al. 2015; Zhang and Seto 2011). Although very-high resolution (VHR) remote sensing imagery (i.e., data with spatial resolution of 1m or less) became more and more available and accessible, no studies at global or very large scales (such as the national scale of China) exist to date. However, several recent studies have demonstrated the feasibility of monitoring the physical expansion and structural urban transformation for smaller test sites (e.g., Huang et al. (2017), Lefebvre and Corpetti (2017), Leichtle et al. (2017a)).

A well and widely known phenomenon with respect to urbanization in China is rural-urban migration, which has persisted for the last decades and was studied intensively in literature (e.g., Liang (2001), Zhang and Song (2003), Chan (2012)). In accordance with the spatial expansion of cities, rural-urban migration was and remains strongly directed towards the coastal provinces in the South-East of China (Chan 2012). Chan (2012) estimated a number of 200-250 million rural residents that have moved to Chinese cities within the past three decades. Despite this continuous and massive process of rural-urban migration in China, a completely different, emerging and less studied phenomenon are so-called *ghost cities.* Briefly, a *ghost city* can be described as a „*new development that is running at severe undercapacity, a place with drastically fewer people and businesses than there is available space for*“ (Shepard 2015). According to Shepard (2015), ghost cities have become a notable phenomenon in China in recent times.

Against this background, this study aims to provide a framework for assessing the ghost city phenomenon by means of multi-temporal VHR remote sensing data. The objective of this study is the assessment of the presence

or absence of the ghost city phenomenon based on the multi-temporal confrontation of available living space of residential buildings (i.e., population capacity) delineated from VHR remote sensing data and multi-temporal population counts from census. For this purpose, a 4d functional city model is established (Section 4.2), comprising a multi-temporal building model, classification of building types associated with residential and non-residential function (i.e., with regard to permanent residential population) as well as building height in terms of the number of floors. Based on these multi-temporal physical morphologic characteristics of the built environment, a statistical approach is applied for multi-temporal population capacity estimation (Section 4.3). The relationship between the multi-temporal evolution of a city's morphologic configuration, building function and related population capacity estimates and the temporal development of its respective census counts enables the assessment whether the ghost city phenomenon is existent or not (Section 4.4). In this study, the city of Dongying serves as an example of a typical highly dynamic medium sized Chinese city, which may suffer from the ghost city phenomenon.

2. Ghost cities of China

The term *ghost city* (or alternatively *ghost town*, *ghost estate*) was originally used for empty or unfinished residential developments in Ireland which occurred as a consequence of the global financial crisis (O'Callaghan et al. 2014). However, the phenomenon of ghost cities became also notable in China recently (Shepard 2015). As opposed to ghost estates in other countries around the globe which mostly arise from abandoned old residential developments (O'Callaghan et al. 2014; Thompson and de Beurs 2018; Xie et al. 2018), a Chinese ghost city is unique since completed new buildings are affected (He et al. 2016). As mentioned above, a *ghost city* in China is defined as a new urban development running at severe undercapacity, a place which houses drastically fewer people and businesses than there is space available (Shepard 2015). A very similar definition is given by Zheng et al. (2017b), who characterize ghost cities by three main features: low population density, low luminosity intensity, and extensive non-lit built-up areas. In another study, Zheng et al. (2017a) define ghost cities likewise as areas of low intensity of anthropogenic activities and dark built-up areas with few lights at night. Jin et al. (2017) associate ghost cities with low urban vitality of new residential developments, which is defined according to urban morphology, functional aspects, and social indicators. Finally, the authors of Ma et al. (2018) relate ghost cities to imbalanced urbanization of land and population as well as low nighttime light intensity.

2.1 Background of the ghost city phenomenon in China

In China, several political, demographic and related spatial policies influence urbanization and population development, which have effects on the evolution of the ghost city phenomenon. Chinese urbanization and migration of rural population towards cities has increased massively since the 1980s and remains high at an unprecedented rate (Ren 2013). However, in contrast to many cities of the Global South, urban development and planning as well as migration of population is controlled by administrative orders and master plans from the Chinese government (Ruibo and Linna 2013). Thus, urban planning in China has become a strategic governmental instrument to relieve the pressure on rapidly growing cities (Shepard 2015) or to promote economic growth of selected cities and provinces (Ren 2013). In addition, the sale of land use rights for urban development has become an important source of revenue and taxes for the local governments, who are consistently reinforcing this approach (Zheng et al. 2014a). However, the population development in China is strongly controlled by the hukou (household registration) system, which assigns a registration status of *urban* or *rural* to each citizen that can only be changed by the government (Ren 2013). Thus, permanent migration from rural to urban areas in the formal sense (i.e., hukou migration) is highly regulated and mostly only possible for the very rich, the highly educated or direct family members of citizens with urban hukou (Chan and Buckingham 2008). Therefore, the number of hukou migrants has been remaining relatively stable at low numbers since the 1990s (Chan 2012). For these reasons, construction of new residential areas has increased much faster than the urban population in some areas, which implies an increasing number of vacant apartments (Zheng et al. 2014a). In turn, such situations depict the starting point for the emergence of the ghost city phenomenon.

2.2 Current state of ghost cities in China

One of the most famous ghost cities in China known from the public media is Ordos Kangbashi in Inner Mongolia, which was established in 2003 after the discovery and development of coal and gas deposits in the beginning of the 21st century (Shepard 2015). Besides this popular example, several other ghost cities are collected by means of rankings and listings from different sources. Since 2014, the Chinese third-party organization Standard Ranking has been publishing an annual list of the top 50 ghost cities identified by means of low population density within the built-up area, setting the threshold to 5500 people / km² (Standard Ranking 2015). In 2015, the real estate portal of Fang Holdings (www.fang.com) published a list of 12 ghost cities in China based on the decline of prices in the local residential real estate market (Fang Ranking 2015). Another list

was published in 2015 by the Chinese internet technology company NetEase (www.163.com), who indicate ghost cities by declining population growth and density accompanied by an increasing growth rate of the urban land area and conclude a concentration of ghost cities in the Northeast of China as well as in coastal areas of Shandong province (NetEase Ranking 2015). Besides their own analyses, the authors of Jin et al. (2017) conducted a keyword search (name of each city + "ghost city" in Chinese) on the Chinese search engine Baidu (www.baidu.com) and created a list of ghost cities in order to assess the public sense for this phenomenon. A complete reprint of these lists and rankings can be found in the appendix section of Jin et al. (2017).

The most common to all of these lists and rankings is that they have only few cities in common. In fact, not a single city was identified as a ghost city by all four assessments. Although the list published by NetEase found some spatial concentration patterns of ghost cities, no general nationwide distribution of ghost cities in China could be identified. Thus, more objective and scientific profound methodological frameworks for the assessment of the ghost city phenomenon are required.

2.3 Scientific assessment of the ghost city phenomenon

In the context of scientific analysis of the ghost city phenomenon, only few and very recent studies exist. The most comprehensive overview of ghost cities in China is given by Shepard (2015), who precisely describes their background, causes, implications, and potential future development from a theoretical perspective. A practical study on ghost cities is presented by Chi et al. (2015), who identify vacant housing areas indicating ghost cities based on cellphone positioning data at national scale in China. Jin et al. (2017) developed a ghost city index for all Chinese cities by multi-temporal evaluation of the vitality of cities based on the number of road junctions, functionality at block level, and location based services (LBS) data. Another study in this context is presented by Zheng et al. (2017b), which is one of only two known studies to date utilizing remote sensing data for the detection of ghost cities at medium to high spatial resolution. In their study, Zheng et al. (2017b) employ remote sensing data from multiple sources (i.e., nighttime lights, land cover, and gridded population data) and propose an index for ghost city identification in the Yangtze River Delta. The second study based on remote sensing data from a similar group of authors (Zheng et al. 2017a) proposes a slightly different ghost city index based on built-up area compared to nighttime light imagery and demonstrates this approach in Northeast China for the detection of ghost cities. In a similar manner, Ma et al. (2018) introduce an improved ghost city index which relates the match-degree of the development of built-up area and population development to nighttime light imagery in the city of Ordos.

These studies show capabilities and limitations of remote sensing data for assessing the ghost city phenomenon; however, none of these approaches employs VHR imagery for detailed assessment at the building level. Furthermore, a detailed discrimination of residential and non-residential function related to building types as well as the vertical dimension of urban areas is neglected throughout. In addition, most of the related work does not consider the multi-temporal aspect of the ghost city phenomenon.

3. Conceptual framework

Summarizing the definitions as well as the findings of existing studies on the ghost city phenomenon in China (Section 2), the most common feature of ghost cities is severe undercapacity with respect to population and businesses within the built-up area while the reliable identification remains challenging. Against this background, the conceptual framework for assessment of the ghost city phenomenon in this study is based on the multi-temporal confrontation of available living space of residential buildings (i.e., population capacity) measured by means of VHR remote sensing and multi-temporal population counts (i.e., permanent residential population) from official census data. Compared to exiting studies, especially the multi-temporal assessment adds robustness to the analyses since uncertainties can be reduced by consideration of temporal trends. Thus, the conceptual definition of the ghost city phenomenon in this study is based on the hypothesis that the available living space massively exceeds the actual population within the study area.

Based on this hypothesis, the deductive methodological framework of this paper is presented in Fig. V-1. For estimation of the multi-temporal residential population capacity, a 4d functional city model is established based on a multitude of input data. This model encompasses three components with respect to the characteristics of the physical built environment: *First*, a multi-temporal building model is developed by means of object-based change detection of individual buildings using VHR remote sensing data. *Second*, building types associated with residential and non-residential function are classified for both time steps. *Third*, building heights from a remote sensing based digital surface model are transformed to number of floors for each building. The multi-temporal population capacity is estimated statistically based on the available living space and opposed to multi-temporal population counts which are collected from official census data. Based on the comparison of the estimated population capacity and actual population counts, the presence or absence of the ghost city phenomenon can be assessed. Especially the dynamics of these numbers increase the validity of this study since the multi-temporal assessment mitigates uncertainties due to consideration of only one snapshot in time. Thus, if the temporal

evolution of the estimated population capacity massively exceeds the change of actual population counts, the presence of the ghost city phenomenon is highly likely according to the definition and hypothesis in this paper.

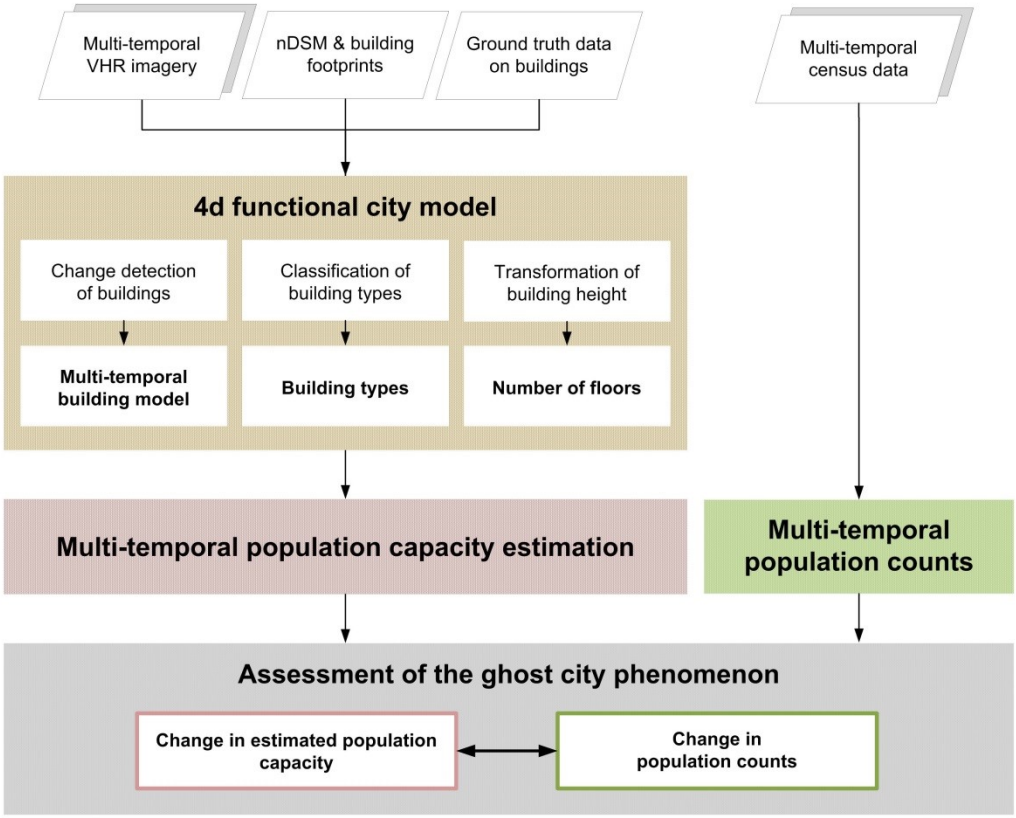

Fig. V-1: The methodological framework of this study.

For demonstration and application of the proposed methodology, a rapidly growing urban area in Shandong province is selected as study area (Fig. V-2). The analysis focuses on the largest and most important city within the area of the Yellow River Delta (YRD), Dongying (Fig. V-2a). After the discovery and development of Shengli Oil Field in the 1960s, the city of Dongying was established in 1983 to harbor oil related industries and its workers (Kuenzer et al. 2014). Since then, the YRD has undergone a rapid economic development accompanied by heavy urbanization and population growth (Ottinger et al. 2013). In Dongying, the local government aims to develop the city as a "national highly-efficient ecological industry base" (Comitee of the Dongying urban and rural planning bureau 2011) and is expediting construction and urbanization. Within its only 30 years of history, the city of Dongying raised to highest recent urbanization rates of the six biggest cities in Shandong province (Xu et al. 2009). In case of Dongying, some studies report the presence of the ghost city phenomenon (e.g., Chi et al. (2015)), whereas other studies do not characterize it as a ghost city (e.g., Standard

Ranking (2015)). Although Dongying is not included in the top 30 ghost cities of China in Jin et al. (2017), it exhibits a relatively high ghost city index according to the authors definition and analyses. Thus, the practical aim of this study is to clarify the presence or absence of the ghost city phenomenon for the city of Dongying as a typical example of a rapidly growing medium sized Chinese city which may suffer from the ghost city phenomenon.

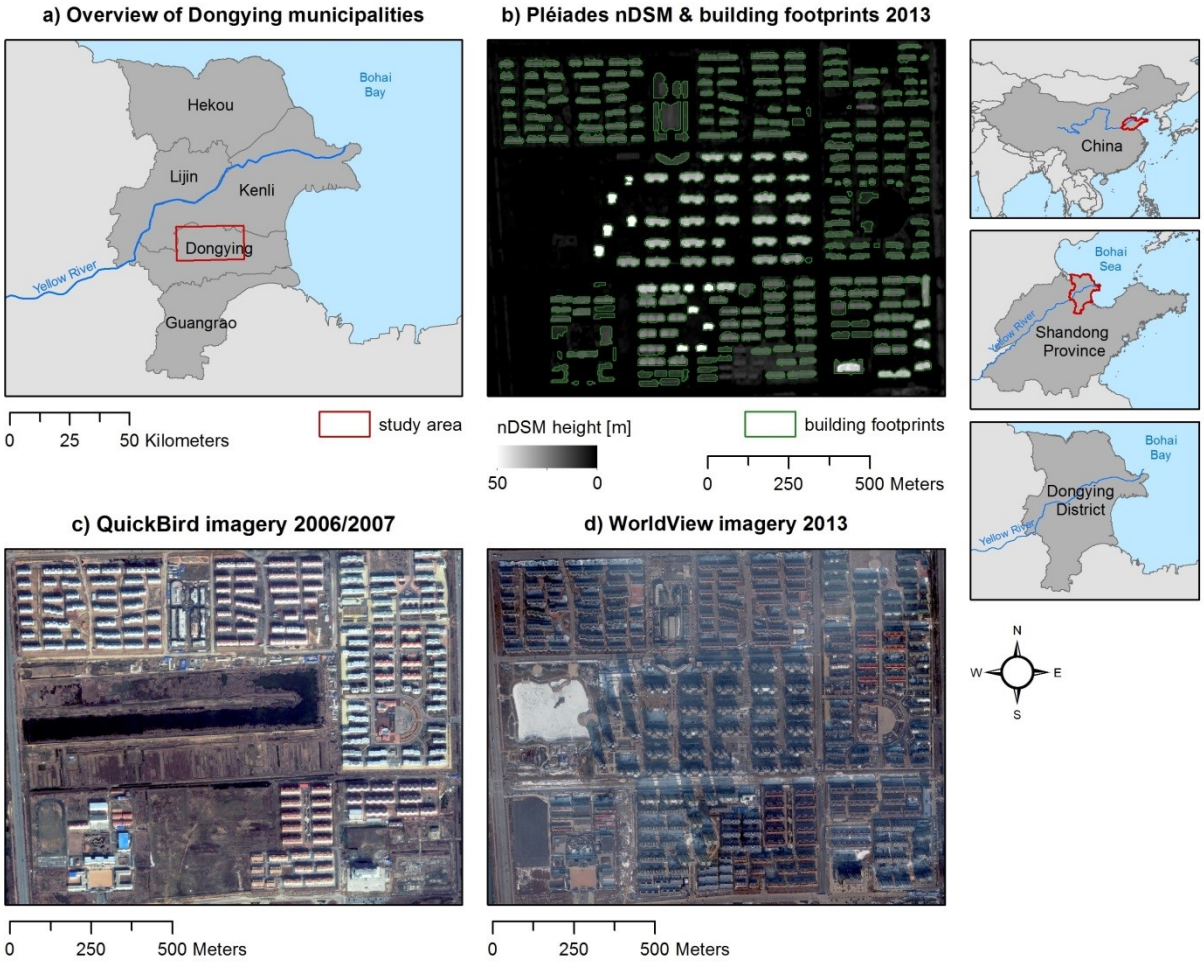

Fig. V-2: Location of the study area and available data sets: a) Overview of the five municipalities in Dongying district and extent of the study area corresponding to the city of Dongying (red frame), b) Zoomed area from the Pléiades normalized digital surface model (nDSM) and derived building footprint geometries, c) Details of the QuickBird imagery from 2006/2007 (t_0), and d) Details from WorldView imagery acquired in 2013 (t_1).

4. Data & methods

4.1 Data sets

4.1.1 Remote sensing data

Multi-temporal VHR remote sensing imagery

The VHR remote sensing imagery serves as the basis for the multi-temporal analysis and consists of two multispectral images with very high spatial resolution. The first image is a mosaic of two QuickBird images which were acquired on March 30, 2006 and February 25, 2007 with off-nadir view angles of 5.7° and 14.2°, respectively (a zoomed area is shown in Fig. V-2c). Based on their close temporal relationship, these images are treated as the first time step 2006/2007 (i.e., t_0). The second image mosaic from 2013 (considered as t_1) consists of three WorldView-2 acquisitions with one image recorded on January 1, 2013 and two images from January 17, 2013 (a zoomed area is shown in Fig. V-2d). The three acquisitions possess off-nadir view angles of 18.1°, 22.5°, and 15.9°, respectively. Both image mosaics exhibit four spectral bands (blue, green, red, and near infrared (NIR)) and were rescaled to a spatial resolution of 50cm using nearest neighbor resampling. The multi-temporal image mosaics were properly co-registered with mean RMSE (root mean square error) of 1.5 pixels (i.e., 0.75m) and atmospherically corrected using the ATCOR (Atmospheric and Topographic Correction) toolbox (Richter 1996).

Normalized digital surface model (nDSM) and building footprint geometries

Building footprint geometries corresponding to the year 2013 (t_1) are available for the urban part of Dongying municipality, i.e., the spatial extent of the city of Dongying (red box in Fig. V-2a). The building footprints were extracted semi-automatically from a normalized digital surface model (nDSM) acquired on October 18, 2013 with spatial resolution of 4m utilizing the segmentation and classification procedure presented in Wurm et al. (2011). The nDSM also provides building height, whereas the median value of all pixels within each building footprint is utilized (also see Tab. V-2). Building heights based on the nDSM possess a coefficient of determination (r^2) of 0.98 compared to building heights from in situ measurements. In addition, the extracted buildings have been checked and corrected manually in order to ensure high reliability (a zoomed area shown in Fig. V-2b). The availability of building footprints delimitates the study area of this paper, which encompasses a total of 30,695 buildings.

4.1.2 Ground truth data on buildings

Ground truth information and reference data on building objects is available for calibration and validation purposes of the 4d functional city model. Changed buildings are defined as newly constructed buildings in this study and reference data on building change was collected via visual inspection of the multi-temporal VHR remote sensing imagery. Accordingly, 7,433 building labels are available for accuracy assessment of change detection which are divided into 3,430 changed and 4,003 unchanged building objects, respectively. Ground truth information on building types associated with residential and non-residential function, which is defined with regard to the occupancy of permanent residential population, as well as height information according to the number of floors is available from field data acquired in October 2013. During this field campaign, the building type and its associated function were determined and the number of floors was counted visually. Within the scope of this study, building functions are reduced to two classes: residential and non-residential, whereas the latter class comprises several classes of building functions (e.g., industrial, commercial, or office buildings). In total, 709 samples on building functions are available, which are separated into 468 residential and 241 non-residential buildings.

4.1.3 Multi-temporal census data

Census data are available from the statistical yearbooks of Shandong province for the past 15 years, i.e., from 2000 until 2015 (Shangdong Provincial Bureau of Statistics 2001-2016), which covers the years that are relevant with respect to the remote sensing data sets (i.e., 2006/2007 (t_0) and 2013 (t_1)) and gives an additional impression of the more long-term trends. This data on the actual permanent residential population is available for all five municipalities in Dongying district. Although the study area largely covers Dongying municipality (Fig. V-2a), the incorporation of the evolution of census counts within the spatial extent of the district area allows for consideration of population development in the wider context of the YRD.

4.2 4d functional city model

4.2.1 Change detection of buildings

In this study, multi-temporal VHR remote sensing data are utilized together with building footprint geometries for object-based change detection in order to derive the multi-temporal building model. The object-based change detection approach is based on the unsupervised methodology presented in Leichtle et al. (2017a).

For the purpose of change detection, building geometries are used to derive object-based features based on the multi-temporal VHR images. Features from different categories, i.e., first-order statistics of the multispectral bands, ratio features, and second-order statistics based on the grayscale image are utilized (Tab. V-1). The object-based features are calculated individually from each VHR image and subsequently combined by means of object-based feature differencing (i.e., difference feature vector) in order to incorporate the temporal relationship implicitly. To ensure equal weighting within the difference feature vector, minimum-maximum normalization is employed to introduce a common range of values of [-1, 1].

Tab. V-1 Object-based features for change detection

Multispectral features (10)	Ratio features (11)	Grayscale texture features (12)
mean intensity of blue band	blue / green	GLCM angular 2nd moment
mean intensity of green band	blue / red	GLCM contrast
mean intensity of red band	blue / NIR	GLCM correlation
mean intensity of NIR band	green / red	GLCM dissimilarity
mean intensity of all bands (brightness)	green / NIR	GLCM entropy
standard deviation of blue band	red / NIR	GLCM homogeneity
standard deviation of green band	normalized green[a]	GLCM mean
standard deviation of red band	normalized red[b]	GLCM standard deviation
standard deviation of NIR band	normalized NIR[c]	GLDV angular 2nd moment
maximum difference between bands	NDVI[d]	GLDV contrast
	SAVI[e]	GLDV entropy
		GLDV mean

grayscale image = (blue + green + red) / 3

[a] normalized green = green / (NIR + red + green), [b] normalized red = red / (NIR + red + green), [c] normalized NIR = NIR / (NIR + red + green)

[d] NDVI = (NIR – red) / (NIR + red), [e] SAVI = [(NIR – red) / (NIR + red + L)] * (1 + L), where L = 0.5

As a final step of feature preprocessing, principal component analysis (PCA) is employed in order to eliminate inevitable correlations and thus redundant information from the object-based difference features in Tab. V-1. For

this purpose, a tripartite procedure is deployed in order to determine the number of relevant principal components (PCs). In this manner, the Bartlett's test is used as a precondition, while the number of relevant PCs is determined by a combined threshold condition of cumulative proportion of explained variance and an averaged parallel analysis (Leichtle et al. (2017a)).

Subsequently, clustering is utilized for discrimination of changed and unchanged buildings based on the relevant PCs of all features. In this paper, four different clustering algorithms that have proven their utility for unsupervised change detection according to Leichtle et al. (2017b-b) are applied for change detection: k-means, partitioning around medoids (PAM), genetic k-means, and the self-organizing map (SOM) clustering approach. Finally, the two partitions of the building inventory are labeled according to changed and unchanged buildings.

The subsequent classification of building types, transformation of height to the number of floors, as well as estimation of population capacity assume that these respective characteristics are stable and do not change within the observation period of 6/7 years. Thus, the information on changed and unchanged buildings is directly incorporated at building level as a precondition for all other components of the 4d functional city model and changed buildings are excluded from the first time step (i.e., t_0).

4.2.2 Classification of building types

Building types are classified based on building geometries and ground truth information on residential and non-residential function. According to Wurm et al. (2016), shape-based features related to building footprints possess high discriminative power for building types classification. The utility of the shape of buildings is also demonstrated in Geiß et al. (2015a) and Xie et al. (2015), where the spatial context of buildings within building blocks is included in terms of additional features. The spatial relationship of buildings is also employed by Lu et al. (2014b) for classification of building types. The current study combines features from these studies, whereas the spatial relationship of buildings is incorporated at the block level, which is defined according to the road network (Tab. V-2).

Tab. V-2 Object-based features for building types classification based on Wurm et al. (2016), Geiß et al. (2015a), Xie et al. (2015), and Lu et al. (2014b).

2d statistics	2d shape features	3d statistics	3d shape features	Spatial relationship of buildings within block
area	compactness	median height	volume / area	mean distance
perimeter	roundness	volume	volume / perimeter	distance to nearest building
area / perimeter	shape index 2d		shape index 3d	mean area
length	asymmetry 2d		asymmetry 3d	SD of area
width	density			mean volume
length / width	rectangular fit			SD of volume
	elliptic fit			mean density
	border index			SD of density
	main direction			mean asymmetry
	normalized perimeter index			SD of asymmetry
	normalized proximity index (nPI)			mean main direction
	normalized spin index (nSI)			SD of main direction
				built-up density

SD = standard deviation

The object-based features according to the shape of buildings and their spatial relationship within each block serve as input for supervised classification of building types using the Random Forest (RF) classifier (Breiman 2001). The required input parameters for the RF classifier comprise the number of classification trees to be generated (n_{tree}) and the number of features that are utilized at each node to establish a decision (m_{try}) (Belgiu and Drăguţ 2016). In the current study, the R software implementation presented by Liaw and Wiener (2002) is employed, where n_{tree} is set to 500 and m_{try} is defined as $\sqrt{p}$, with p corresponding to the number of features. These values are initialized according to default values, which is also recommended by Belgiu and Drăguţ (2016).

Prior to training of the RF classifier, the features are scaled to a common range of values of [0, 1] in order to ensure equal weighting of all features. The ground truth information on building function is available for 709 sample buildings and separated into training and testing data picking a stratified random sample with partitioning of 85% training and 15% testing samples. Subsequently, the RF model is trained and applied to the complete set of buildings for classification of building types according to residential and non-residential function.

4.2.3 Transformation of building height

Since the statistical approach for population capacity estimation is based on the available living space per residential building in this study, building height must be transformed to number of floors for all buildings first (Fig. V-3). For this purpose, the average floor height of buildings with residential function from the reference data set (Section 4.1.2) $h_{fl_j}(o_{b_j} = residential)$ is calculated by linear transformation of building height from the nDSM. In this regard, the height of each reference building (h_{b_j}) is divided by its number of floors (n_{fl_j}), which is available through the nDSM and ground truth information (Section 4.1.2). The average floor height of the reference buildings is compared to the height of all buildings (i.e., their median nDSM value ($h_{b_i}(t_{b_i} = residential)$)) to derive the number of floors (n_{fl_i}) for all residential buildings within the study area. This step includes three threshold conditions based on building height and the statistics of the floor heights from the reference data (Fig. V-3). The first condition excludes buildings with smaller height than the lowest building in the reference data ($n_{fl_i} = 0$). The second condition assigns one floor ($n_{fl_i} = 1$) to buildings that are lower than the average floor height but not included in the first condition. For buildings from the third condition (i.e., buildings with greater height than the average floor height from reference data), their height is divided by the average floor height for residential buildings to determine the number of floors (n_{fl_i}). In this context, only full floors are considered, i.e., the number of floors (n_{fl_i}) is generally rounded down.

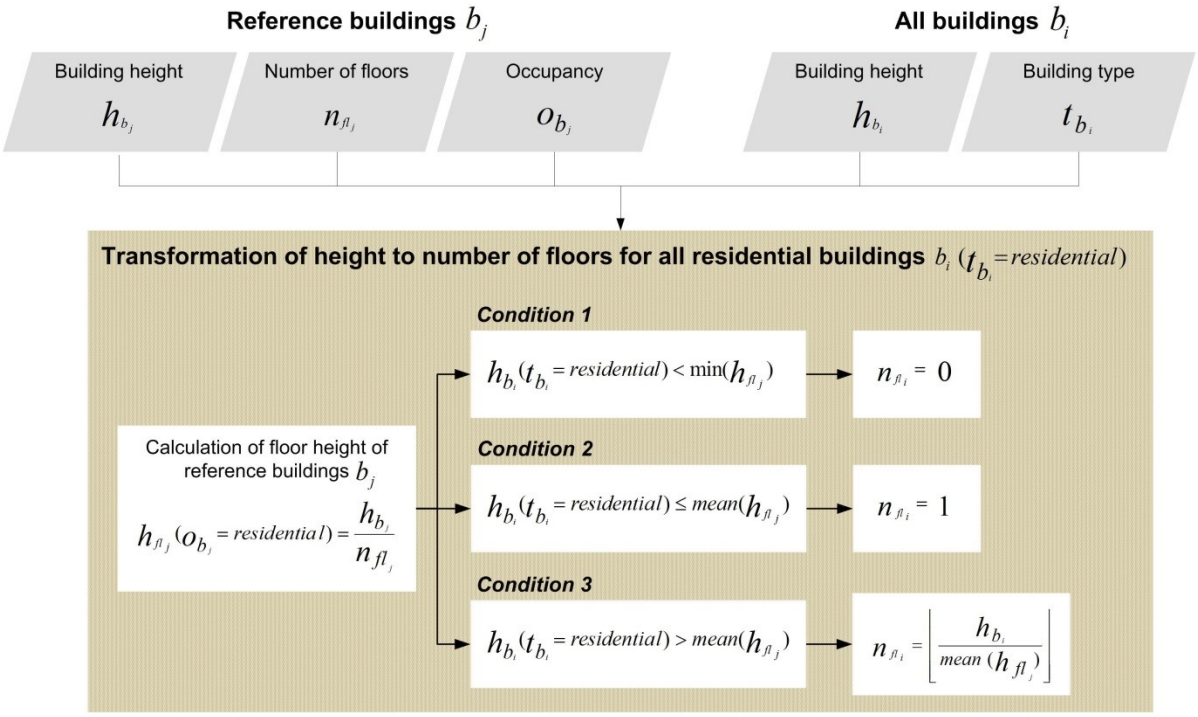

Fig. V-3 Transformation of building height to number of floors for all buildings within the study area. Formulae index j refers to buildings from the reference data set, index i indicates the complete building inventory within the study area. For mathematical symbols and abbreviations see text.

4.2.4 Validation of the 4d functional city model

For validation of the 4d functional city model, each parameter (i.e., the multi-temporal building model, building types classification, and number of floors) is validated separately according to its respective ground truth information. For accuracy assessment of the multi-temporal building model as well as building types classification, the Kappa statistics (Foody 2002), the F-score (Van Rijsbergen 1979), as well as the True-Skill-Statistics (TSS) (Allouche et al. 2006) are calculated based on the respective ground truth information (Section 4.1.2). The linear transformation of building height to the number of floors is validated based on the reverse prediction of the number of floors using the average height per floor of reference buildings with residential function, whereas the coefficient of determination (r^2) is calculated as a quantitative measure of accuracy.

4.3 Population capacity estimation

Population capacity at the building level is estimated according to the statistical approach for population estimation (Biljecki et al. 2016). The combination of the classification of building types and the area of building footprints together with the number of floors allows the spatial quantification of available living space. Fig. V-4 presents the workflow for population capacity estimation. First, the available living space for each residential building ($ls_i(t_{b_i} = residential)$) is estimated by multiplication of the number of floors (n_{fl_i}) with the area of each floor based on the building area (a_{b_i}). Second, the population capacity of each building ($pop_{capacity_i}$) is estimated based on the available living space (ls_i) and statistical data on living space per capita (ls_{capita}) available from the Comitee of the Dongying urban and rural planning bureau (2011). The estimated population capacity for each building is generally rounded down. Based on the multi-temporal building model, multi-temporal population capacity estimates are established.

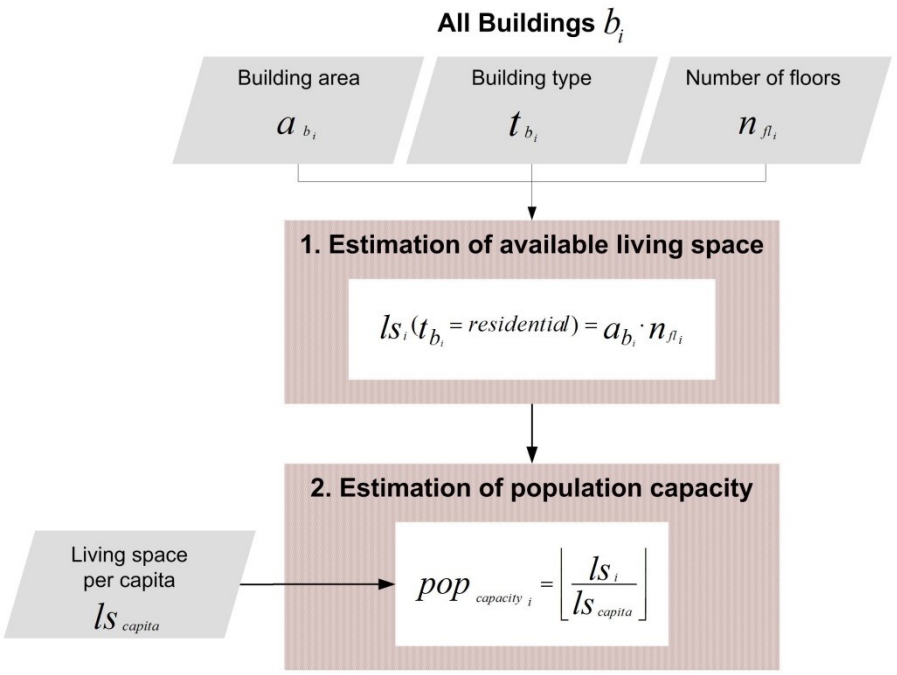

Fig. V-4 Workflow for estimation of population capacity. For mathematical symbols and abbreviations see text.

4.4 Assessment of the ghost city phenomenon

According to the conceptual framework of this study (Section 3), the presence or absence of the ghost city phenomenon is assessed by the comparison of multi-temporal population capacity estimates (Section 4.3) against multi-temporal census counts (Section 4.1.3). The estimated numbers of population capacity per building are summarized for the complete study area in order to approximate the administrative municipality level of census data. Accordingly, if the estimated population capacity massively exceeds population counts, the presence of the ghost city phenomenon is highly likely. In contrast, if these numbers correspond with each other or exhibit only slight variation, there is no indication for the presence of the ghost city phenomenon. In order to provide a more robust assessment, the dynamics of population capacity are opposed to the change of population counts in terms of multi-temporal assessment.

In addition, this comparison is also conducted in a spatially more detailed manner to provide an impression of specific regions within the urban area which are most likely to suffer from the ghost city phenomenon. According to the conceptual framework, a higher probability for the presence of the ghost city phenomenon can be assumed in areas of most recent residential developments based on constant or decreasing population numbers according to the census data. For the detailed spatial localization, the dynamics of the estimated population capacity (epc) and census counts (cc) are scaled to relative values of [-1, 1] due to the deviation of spatial scales and consequent deviation of absolute numbers of both parameters (i.e., epc at building level, cc on administrative units):

$$\rho\ epc = \frac{\Delta\ epc}{|\max(\Delta\ epc)|}$$

$$\rho\ cc = \frac{\Delta\ cc}{|\max(\Delta\ cc)|}$$

with $\Delta\ epc = epc\ (t_1) - epc\ (t_0)$

and $\Delta\ cc = cc\ (t_1) - cc\ (t_0)$.

The dynamics of estimated population capacity and census counts are combined in terms of an index as follows:

$$\frac{\rho\ epc - \rho\ cc}{2}$$

This index indicates the relationship of population capacity and actual number of residents with respect to the smallest spatial scale of inputs (i.e., the building level in this study) and possesses a range of values of [-1, 1].

Accordingly, values towards 1 represent high likelihood for the presence of the ghost city phenomenon, while values around 0 refer to a proper degree of occupancy and low probability for the presence of a ghost city. Increasing census counts with simultaneous decreasing availability of living space (i.e., population capacity) result in negative values and would indicate housing shortage, i.e., the reverse situation of the ghost city phenomenon.

5. Results

5.1 4d functional city model

5.1.1 Multi-temporal building model

For development of the multi-temporal building model and observation of the physical growth of the city of Dongying, the approach for unsupervised change detection was performed using different clustering algorithms on the WorldView-2 and QuickBird images from 2013 and 2006/2007, respectively. The change detection results of the four clustering algorithms (i.e., k-means, PAM, genetic k-means, and the SOM clustering approach) were compared according to their classification accuracy and the respective best result was used (Fig. V-5a).

In general, Dongying is a highly dynamic city, which is clearly reflected by the results of change detection analysis. Within the considered time span of 6/7 years, 12,875 new buildings were constructed, which corresponds to an increase of the building inventory of 72%. As shown in Fig. V-5a, new buildings were built more or less equally distributed within the city area of Dongying. Notably, complete blocks were established in most cases, whereas construction of single buildings represents an exceptional case.

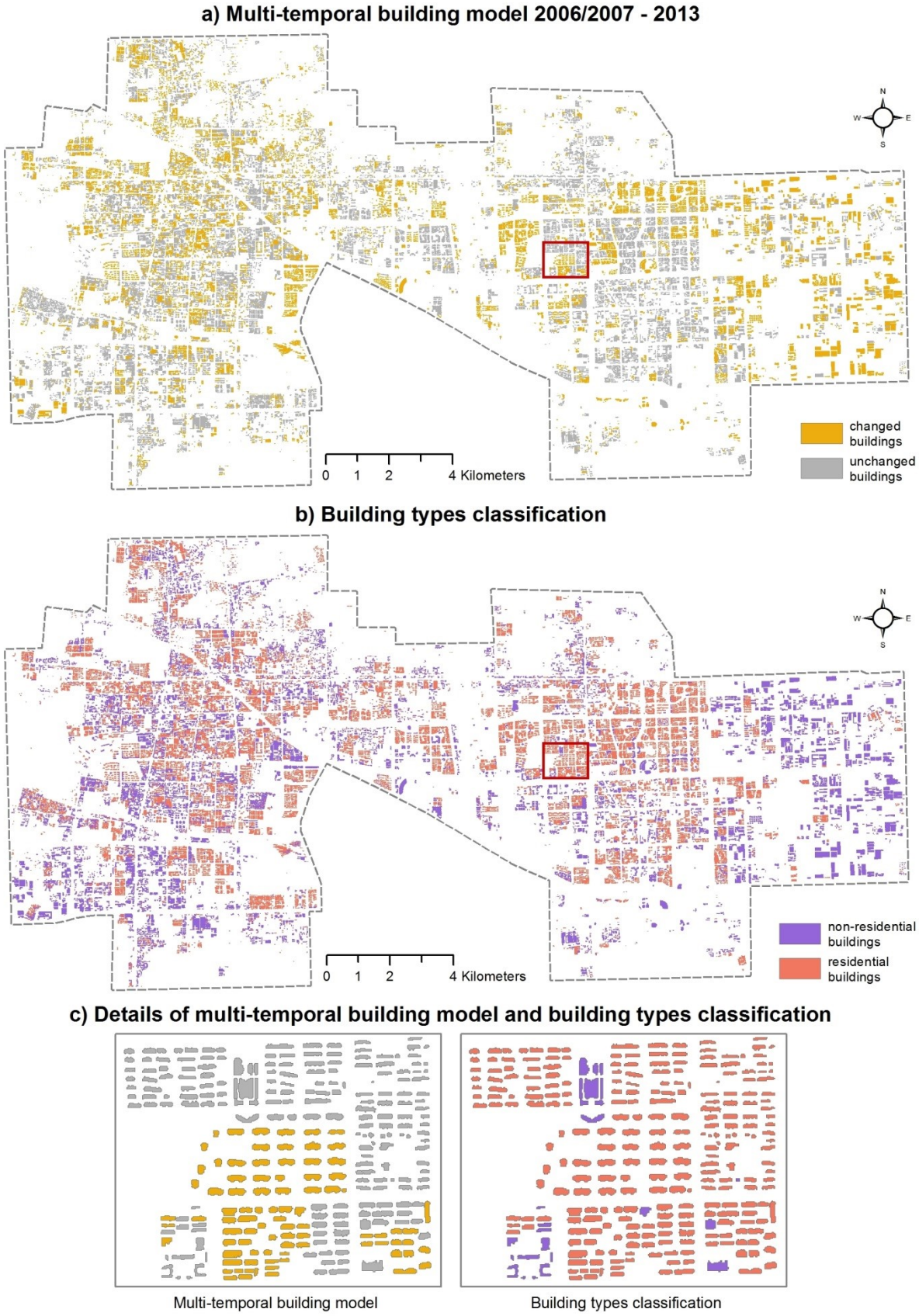

Fig. V-5 Change detection and classification results in Dongying. a) Multi-temporal building model according to unsupervised change detection of QuickBird imagery from 2006/2007 (t_0) and

WorldView-2 acquisitions from 2013 (t_1). b) Building types classification associated with residential and non-residential function. c) Details of the multi-temporal building model and building types classification.

5.1.2 Building types

The supervised RF classification of building types was conducted based on object-based features at building level (Tab. V-2) and ground truth information according to residential and non-residential function (Section 4.1.2). After scaling of all features to a common range of values of [0, 1], the ground truth data were stratified and randomly sampled into 398 + 205 training and 70 + 36 testing samples for the two classes of interest (i.e., building types with respect to residential and non-residential function), respectively.

Fig. V-5b depicts the classification result of building types associated with residential and non-residential function. In total, 18,166 buildings were classified as residential and 11,821 were identified as non-residential. As shown in Fig. V-5b, the majority of residential buildings are concentrated in the central parts of the city of Dongying, while non-residential buildings are located more towards the outskirts. Especially the Eastern part of the city is dominated by large non-residential buildings which can be allocated industrial function. Furthermore, building types are much more diverse in the Western part of Dongying, i.e., residential and non-residential buildings are allocated in a mixed manner at small spatial scales. In relation to the multi-temporal building model (Fig. V-5a), 8,351 (65%) newly constructed buildings correspond to residential function while 4,524 (35%) account for non-residential buildings.

5.1.3 Number of floors

For the purpose of transformation of height from the nDSM to the number of floors for all residential buildings within the study area, the average height per floor was calculated using the relationship of building height and the number of floors for all reference buildings with residential function. Fig. V-6 depicts the relationship between building height and the number of floors of the reference buildings which serves as the basis for linear transformation of building heights from the nDSM. In this study, an average height per floor of 2.88m was empirically derived based on 468 buildings with residential function from the ground truth data set. Based on the average height per floor, the number of floors was calculated using the median height from the nDSM for all residential buildings within Dongying.

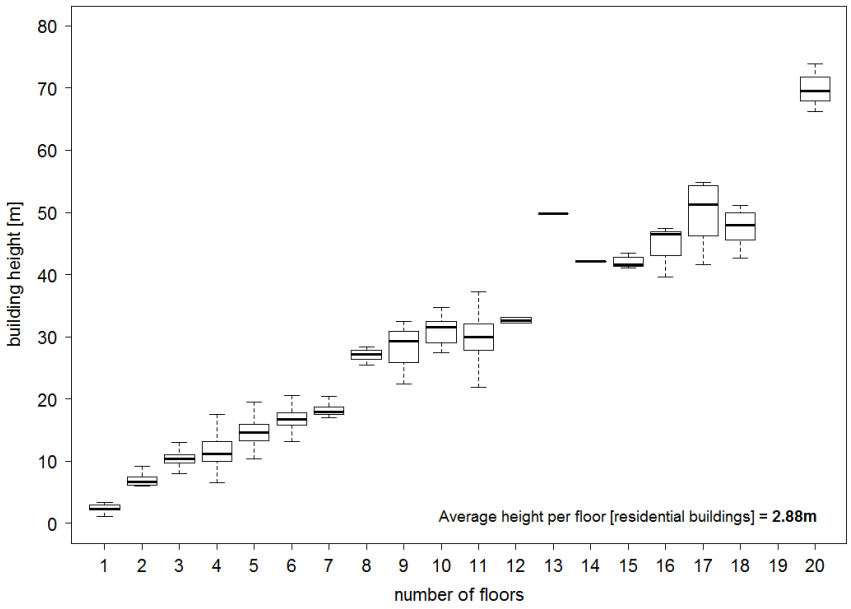

Fig. V-6 Relationship of building height (i.e., median height from the nDSM per building footprint) and number of floors for 468 reference residential buildings within the study area.

5.1.4 Validation of the 4d functional city model

For validation of the multi-temporal building model (Fig. V-5a), the change detection accuracies of the four clustering algorithms were merged and returned final weighted accuracies of $\kappa = 0.73$, $F = 0.85$, and $TSS = 0.73$ for the complete city extent. With respect to building types classification (Fig. V-5b), the testing samples of the RF classification were used for validation with accuracies of $\kappa = 0.89$, $F = 0.93$, and $TSS = 0.89$. The linear transformation of building height to the number of floors was validated according to the reverse prediction of the number of floors with a coefficient of determination of 0.91.

5.2 Population capacity estimation

The 4d functional city model serves as input for multi-temporal population capacity estimation according to Fig. V-4. The population capacity was calculated based on the available living space per building and the statistical living space per capita, which is 35 square meters per person for the city of Dongying according to the Comitee of the Dongying urban and rural planning bureau (2011). Fig. V-7a shows the estimated population capacity for all residential buildings in Dongying for the year 2013. The population capacity for 2006/2007 (Fig. V-7b) is given implicitly based on the multi-temporal building model.

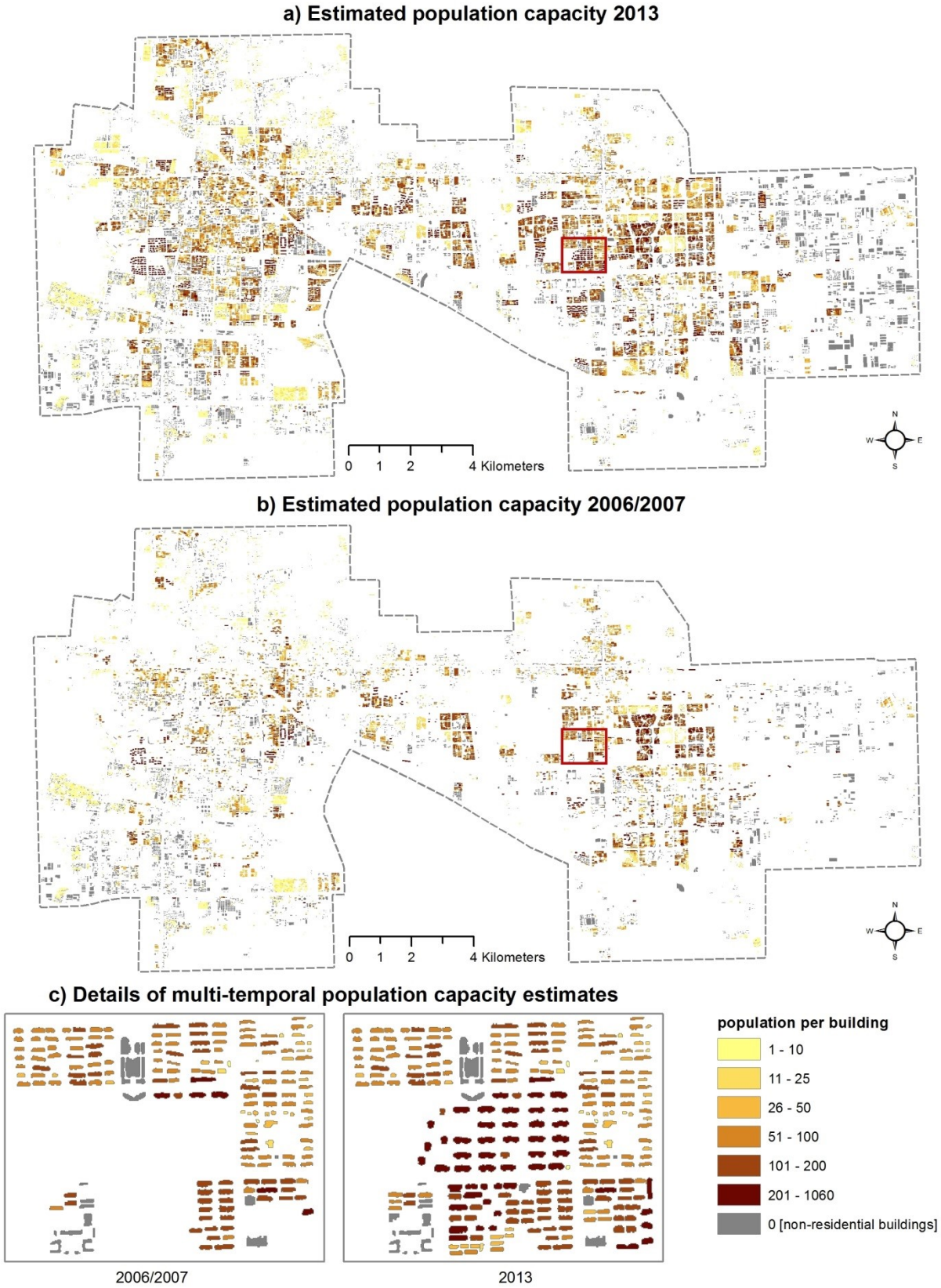

Fig. V-7 Multi-temporal estimation of population capacity at building level for the city of Dongying. a) Population per building in 2013 (t_1). b) Estimated population capacity in 2006/2007 (t_0). c) Details of multi-temporal population capacity estimation.

As shown in Fig. V-7, the general spatial pattern of the estimated population capacities reflects the combined results of change detection and building types classification, while large high-rise residential buildings provide living space for a large number of people as opposed to low buildings with small area. The smallest buildings in Dongying provide room for one person each, while the largest residential building is capable of hosting 1,060 people. In total, the summarized population capacity of Dongying is 587,914 people for 2006/2007, while this number increases to 1,188,749 for the year 2013. The increase of the building inventory corresponds to an increase of population capacity of 102% within the considered time span of 6/7 years.

5.3 Assessment of the ghost city phenomenon

For the assessment of the presence or absence of the ghost city phenomenon, the aggregated numbers of estimated population capacity in Dongying were related to census counts of the administrative unit of Dongying municipality. Although the study area possesses a large spatial overlap with Dongying municipality (Fig. V-2a), census counts of all five municipalities (i.e., Dongying district) were consulted in addition for consideration of the wider context of the YRD. The multi-temporal comparison of estimated population capacity and census counts (Fig. V-8) according to the statistical yearbooks for Dongying municipality (blue graph in Fig. V-8) show very good agreement in 2006/2007 (estimated total population capacity of 587,914 opposed to 611,000 people in 2006 and 616,000 people in 2007 according to census), but large disagreement in 2013 (estimated total population capacity of 1,188,749 opposed to 598,000 people according to census). Thus, the estimated population capacity in 2006/2007 deviates 4.17% from the census counts, while in 2013 the relative difference is 98.79%. The development of census counts in Dongying district according to the statistical yearbooks (orange graph in Fig. V-8) underlines the validity of the evolution of population counts in Dongying municipality. In conclusion, the estimated population capacity massively exceeds population counts in 2013 and thus, the assessment in this study reveals severe undercapacity with respect to permanent residential population within the urban area of Dongying. In other words, the emergence of the ghost city phenomenon can be considered highly likely within the time period of monitoring according to the definition of this study.

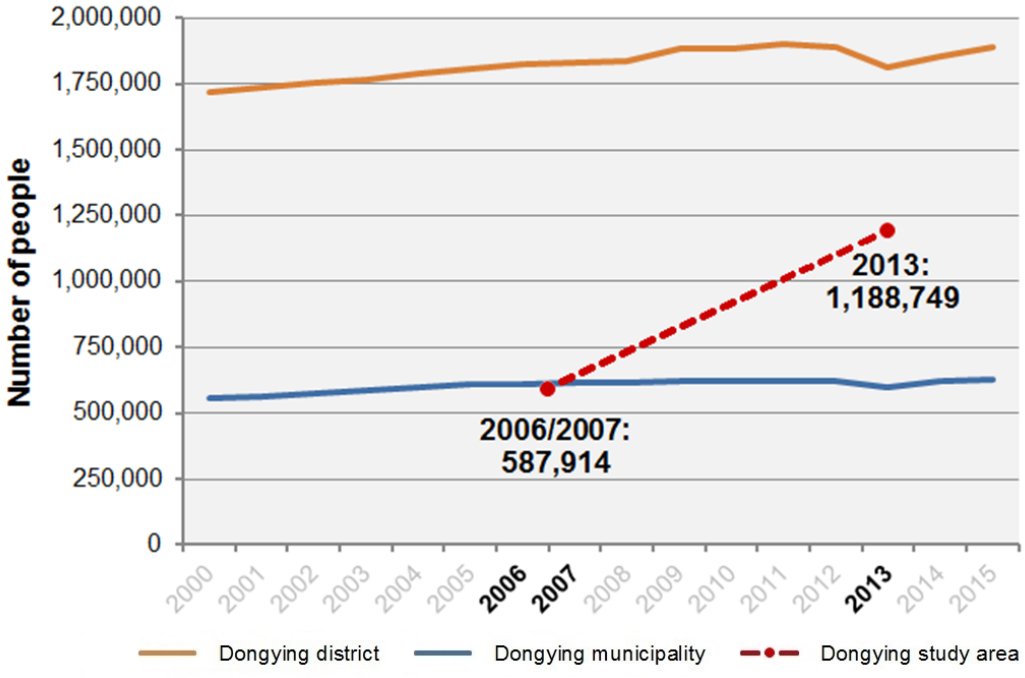

Fig. V-8 Multi-temporal comparison of population capacity estimates in 2006/2007 (t_0) and 2013 (t_1) (red) with census counts of Dongying district (orange) and Dongying municipality (blue).

The calculated index describing the ghost city phenomenon and its spatial distribution is presented in Fig. V-9. Besides index values of individual buildings, its mean value at the block level (defined according to the road network) is displayed for spatial aggregation and generalization and reveals specific regions of the urban area which are most likely to suffer from the ghost city phenomenon. Obviously, these areas correspond to the synopsis of the 4d functional city model and multi-temporal population capacity estimates. Based on the city-wide findings that in 2006/2007 the estimated population capacity matches census counts quite exactly, specific spatial conclusions are feasible. Since census counts are only available for the complete city extent, an equal trend is implicitly applied to all residential buildings in this study. However, assuming that the majority of the residents in Dongying do not migrate within the city during the 6/7 years of monitoring and the actual population of the city does not change, newly established quarters of large and high-rise residential buildings are most likely to suffer from the ghost city phenomenon (Fig. V-9), which was also concluded by Shepard (2015), Jin et al. (2017), and Ma et al. (2018). Accordingly based on the more or less stable population of Dongying city, existing (i.e., unchanged) residential areas are less likely to suffer from the ghost city phenomenon as shown in Fig. V-9.

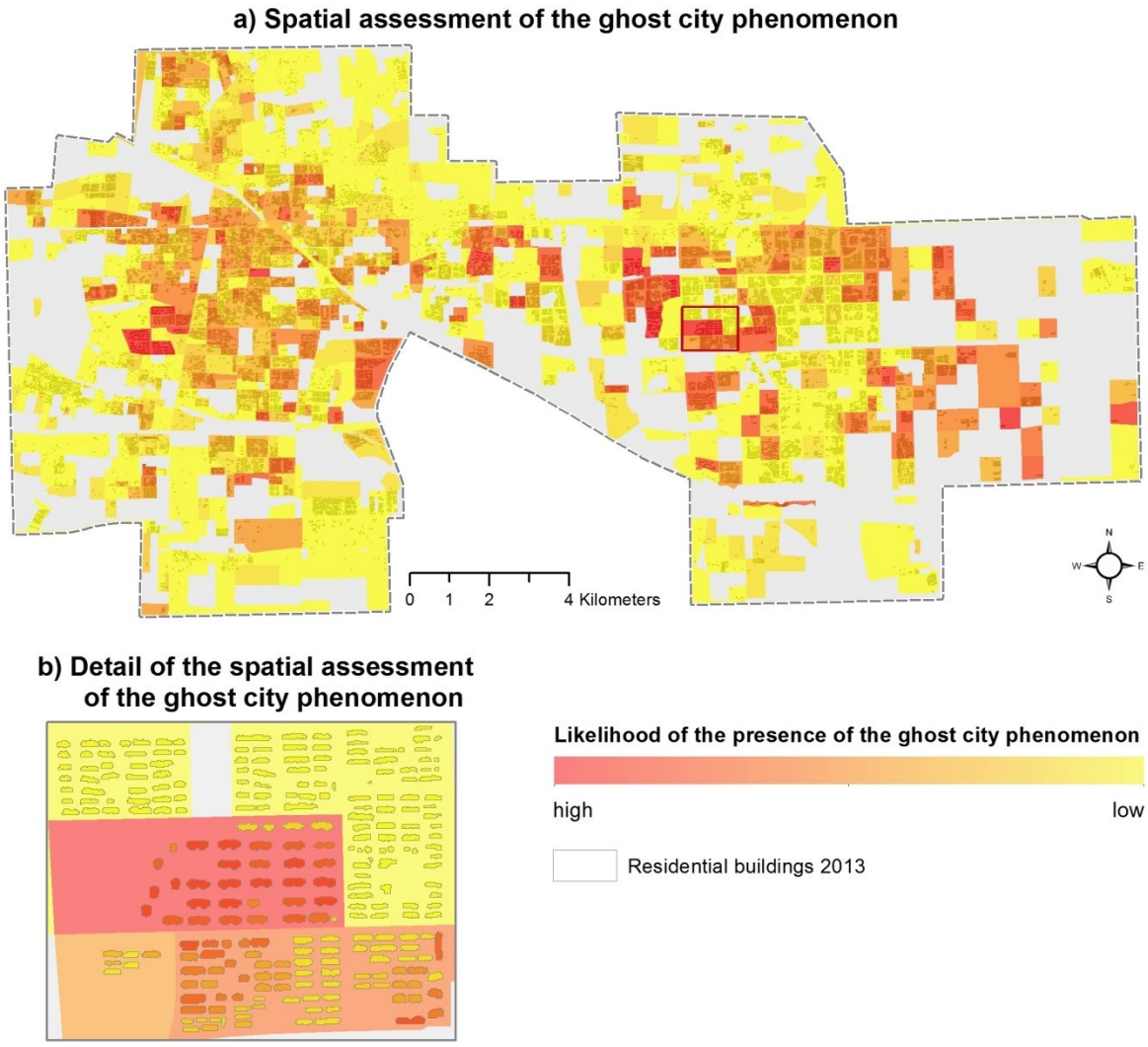

Fig. V-9 Spatial dynamics of the estimated population capacity related to census counts: a) Spatial assessment of the ghost city phenomenon, and b) Detail of the spatial assessment of the ghost city phenomenon.

6. Discussion

This paper demonstrated a framework for the assessment of the ghost city phenomenon based on VHR remote sensing data as well as census counts and revealed significant evidence for the emergence and presence of the ghost city phenomenon in case of the Chinese city of Dongying.

Foremost, the proposed methodological framework is based on a 4d functional city model. The methodology of unsupervised change detection for setup of the multi-temporal building model is discussed in detail in Leichtle

et al. (2017a), whereas the obtained accuracy for the study area of $\kappa = 0.73$, $F = 0.85$, and $TSS = 0.73$ is similar to related work (e.g., Huang et al. (2017), Lefebvre and Corpetti (2017)). With respect to the classification of building types, numerous studies have proven the utility of the RF classifier since its introduction by Breiman (2001). The accuracy of building types classification ($\kappa = 0.89$, $F = 0.93$, $TSS = 0.89$) was comparable to existing studies in literature (e.g., Belgiu et al. (2014), Geiß et al. (2015a), Wurm et al. (2016)). The number of floors was determined based on linear transformation of building height with a high coefficient of determination of 0.91. Thus, the 4d functional city model provides a reliable data set of high thematic and geometric detail as well as high accuracies for characterization of the physical morphology and building functionality of the city. Subsequently, this information was employed for population capacity estimation based on a statistical approach for city-wide assessment of available living space and the potential number of inhabitants at building level. Other studies in literature found good accuracies using the statistical approach for population estimation, e.g., Wang et al. (2016b) achieve a mean relative error of 16.46% compared to census counts while Tomás et al. (2016) underestimate the population by only 1.36%. In comparison, the estimated population capacity in 2006/2007 from this study deviates 4.17% from the respective census count for Dongying municipality, which indicates the validity of the relationship between the 4d functional city model with estimated population capacity and census data. However, quantitative validation of the multi-temporal population capacity estimation is challenging since detailed area-wide, reliable ground truth information is not existent in this regard.

In addition, the methodological framework in this study encompasses known sources of error. *First*, the accuracy of change detection could be improved slightly by adjustment or correction of building parallax distortions, especially in case of high-rise buildings as presented in Leichtle et al. (2017a). Furthermore, buildings under construction are classified as changed buildings and are included in the population estimation. However, as construction time is comparatively short in China, this leads to a short temporal inconsistency only. *Second*, mixed functions within individual buildings (e.g., commercial use on the ground-floor of residential buildings along shopping streets) might blur the precision with regard to classification of building types. In addition, the linkage of building types and associated functions might introduce uncertainties since it may not be unambiguous for all buildings. *Third,* the average floor height was calculated based on linear transformation of height for all residential buildings that includes certain variation (also see Fig. V-6), which is not represented by the linear model. *Fourth*, errors and uncertainties with respect to the geometry of building footprints lead to loose representation of features for change detection as well as for classification of building types, but also cause under- or overestimation of available living space, especially in case of high-rise residential buildings.

Nevertheless, the accuracy and the depth of detail in the 4d functional city model surpass existing studies and provide a reliable description of the physical morphologic configuration and associated function within the city. *Finally*, the estimation of population capacity is dependent on several factors: Although the classification of building types is of good accuracy, there might be residential buildings of mixed functions as described above or additional non-residential space within residential units (e.g., storage space, utility rooms, garages, etc.) might exist that is not accounted to living space in reality. In this context, as mentioned above, the linear estimation of the average floor height influences the accuracy of the calculation of the number of floors. Furthermore, the living space per capita may vary considerably dependent on social and economic factors within the city area (Wu et al. 2008), e.g., villa-type residential buildings possess significantly higher living space per capita compared to residential buildings of ordinary or poor living conditions. In addition, the utilized value of living space per capita is based on the year 2013 as a reference and could have changed (most likely increased) since 2006/2007 which might introduce a mismatch (most likely an overestimation) of population capacity. The utilized value of living space per capita is of high importance in general, since it significantly determines the absolute numbers of population estimation, as highlighted by Taubenböck and Wurm (2015). Nevertheless, this uncertainty is anticipated by the multi-temporal population estimation since the temporal characteristics of population capacity estimates are preserved regardless of the exact number of living space per capita.

A crucial point for the comparison of population capacity estimates against census counts are uncertainties related to census data from the statistical yearbooks employed in this study. The mismatch of administrative boundaries and the spatial extent of the remote sensing data analysis induces sources of error. Due to the large overlap of the study area with Dongying municipality (Fig. V-2a), census counts of this administrative unit were considered in particular while the numbers of all five municipalities (i.e., Dongying district) were consulted additionally for incorporation of the wider context of the YRD. Besides the general uncertainties in the inference of census counts on the permanent residential population (Simpson 2007), an important peculiarity of census counts in China is the hukou (household registration) system, which assigns a registration status of ‘urban’ or ‘rural’ to each citizen that is reflected in census numbers and can only be changed by the government (Ren 2013). Thus, permanent migration from rural to urban areas in the formal sense (i.e., hukou migration) is highly regulated and mostly only possible for the very rich, the highly educated or direct family members of citizens with urban hukou (Chan and Buckingham 2008). In fact, most of the rural-urban migrants are rural migrant workers (i.e., people with rural hukou working in urban areas), who are not captured by official census counts and whose number increased to 221 million people in 2010 according to Chan (2012), whereas 7 million people account for Shandong province (total population of Shandong province according to census in 2010: 95,790,000)

(Liang et al. 2014). At the same time, the number of hukou migrants (i.e., permanent migrants) remained relatively stable at low numbers (Chan 2012). This relatively high number of floating population in turn also affects census counts which represent de jure population (in contrast to de facto population) (Wu et al. 2005). However, rural migrant workers have very limited access to the local real estate market due to their rural hukou status and mostly live in dormitories provided by their employers (Wong et al. 2007).

For the decision on the presence or absence of the ghost city phenomenon, the dynamics of estimated population capacity are opposed to the dynamics of census counts. With respect to the comparison at the city level, the estimated population capacity matches census counts in 2006/2007 very well but a large disagreement was found in 2013 for the example of the city of Dongying. Nevertheless, both numbers are not free of errors which could imply either an underestimation of population by means of census or an overestimation of population capacity justifying the mismatch in 2013. However, the temporal evolution of economic indicators in Dongying municipality and Dongying district (Shangdong Provincial Bureau of Statistics 2001-2016) clearly underlines the strong economic as well as the rapid urban development. For example, according to the statistical yearbooks of Shandong province, the Gross Domestic Product (GDP) increased by 109% in Dongying district (GDP in 20006/2007 155,756 Mio. ¥, GDP in 2013 325,020 Mio. ¥) while the investment in fixed asset increased by 252% in Dongying municipality (12,089 Mio. ¥ in 2006/2007, 42,539 Mio. ¥ in 2013). These numbers support the strong construction activity identified by the 4d functional city model since expansion of urban land is highly correlated to growth of GDP in China (Seto et al. 2011). In addition to the quantitative accuracy assessment of the multi-temporal building model, this underlines the validity of the dynamics of population capacity estimation. In contrast, the accuracy of census counts remains unclear as mentioned above since reliable and detailed ground truth information is not existent in this regard. Regardless of these uncertainties, this study revealed a viable large multi-temporal disagreement of the change in available living space (i.e., population capacity) compared to the temporal development of census counts (Fig. V-8) which confirms the presence of the ghost city phenomenon. The detailed spatial assessment (Fig. V-9) provides an impression of specific regions of the urban area which are most likely to suffer from the ghost city phenomenon and revealed that large new residential developments are most likely affected. Since the dynamics of census counts are only available at the administrative municipality level which covers the complete study area, the particular spatial variability of the permanent residential population distribution at the building level cannot be captured due to the lack of more detailed data. However, the detailed spatial results appear feasible based on the conceptual definition of the ghost city phenomenon in this study and also coincide with the findings of Shepard (2015), Jin et al. (2017), and Ma et al. (2018). In addition, the methodology of the proposed index allows

utilization of more detailed census counts for a more accurate spatial assessment if such data is available, however this might shift the arrangement of the detailed spatial assessment due to issues with respect to modified areal units (Openshaw 1984).

In view of existing studies assessing the ghost city phenomenon, Standard Ranking (2015) did not identify Dongying being part of the top 50 ghost cities of China. In contrast, Chi et al. (2015) include Dongying in their list of 20 Chinese ghost cities based on GPS data. Although Jin et al. (2017) do not incorporate Dongying in their final ranking of the top 30 ghost cities of China, it is evident from their analyses that Dongying exhibits a relatively high ghost city index according to the authors definition. As described above, this study found feasible evidence for the emergence and presence of the ghost city phenomenon in the city of Dongying. Nevertheless, the future development of construction activities as well as census counts must be observed and analyzed according to the methodological framework presented in this study in order to provide continuous monitoring of the evolution of the ghost city phenomenon in Dongying.

7. Conclusion

This paper presents a framework for multi-temporal and spatially detailed assessment of the presence or absence of the ghost city phenomenon by means of multi-temporal VHR remote sensing data using the Chinese city of Dongying as an example. The conceptual framework of this study is based on the hypothesis that the available living space massively exceeds the actual population in case of presence of the ghost city phenomenon. In detail, the presence or absence of the ghost city phenomenon is assessed by multi-temporal confrontation of available living space of residential buildings (i.e., population capacity) and multi-temporal census counts. The multi-temporal population capacity estimation based on a 4d functional city model incorporates a multi-temporal building model, building types in terms of function, as well as the number of floors of each building. In case of Dongying, the multi-temporal analysis revealed a considerable mismatch of estimated population capacity based on remote sensing and permanent residential population numbers from census counts. In other words, the physical manifestation of urbanization does not match the population development and thus, this study found reasonable evidence for the presence of the ghost city phenomenon in Dongying.

In general, the ghost city phenomenon mostly affects rapid developing small and medium sized cities in China (Jin et al. 2017). In China, a crucial factor with respect to this phenomenon is the strong influence of planning on urban development, which is largely controlled by the government. From a more general perspective, Shepard

(2015) argues that in fact, the ghost city phenomenon represents a transition period of construction and occupancy of newly constructed buildings, which is a common process in newly developed districts in Chinese cities. The open question of all new developments is whether and when these new urban centers will be populated, which still remains unanswered for the city of Dongying.

Thus, continuous monitoring is indispensable for present ghost cities like Dongying in order to reach a final decision on the persistent presence of the ghost city phenomenon and their future development. From a data perspective, additional information especially on human activities might be included as additional indicators for comprehensive understanding of the characteristics and evolution of potential ghost cities. Consequently, the employed methodology must be constantly improved and emerging techniques for supervised as well as unsupervised image classification and change detection as well as associated object-based features must be evaluated and included. For broader understanding of the ghost city phenomenon and related processes, future investigations must also incorporate additional Chinese cities and larger spatial scales.

Acknowledgements

This research was conducted in the context of the DELIGHT project (www.delight.eoc.dlr.de, grant no. 02WCL1249A and 02WCL1249I) funded by the German Federal Ministry of Education and Research, BMBF. Additional funds were received from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant no. 714087 - So2Sat). The authors would like to thank Matthias Boes (European Space Imaging, EUSI) for providing the WorldView imagery and Dr. Klaus Martin (Company for Remote Sensing and Environmental Research, SLU) for his support. The authors would also like to thank the anonymous reviewers for their very helpful comments and suggestions.

Chapter VI. Synthesis

1. Summary and discussion

The main goal of this dissertation is the detection of changes in an urban environment based on multi-temporal VHR remote sensing imagery from different sensors and consequent exemplary application in the context of urban geography. To meet this aim, a suitable object-based approach for change detection of individual buildings in VHR data was developed and subsequently applied for the assessment of the ghost city phenomenon in a typical highly dynamic city in China. In detail, this thesis targets the following three specific objectives that contribute to the overall aim and fill existing research gaps.

1) Development of a change detection approach for VHR data from different sensors

The change detection approach in this thesis is implemented according to the object-based paradigm, focusing on the construction of individual buildings, which represents one of the most distinctive morphological changes in the context of urban growth. Technically, as presented in Chapter II, the workflow for change detection includes comprehensive preparation of object-based difference features which employs principal component analysis (PCA) together with a unique procedure for determination of the number of relevant principal components followed by k-means clustering for discrimination of changed and unchanged buildings. Dependent on the utilized set of object-based input features, the proposed approach returned viable results in the order of 0.8 to 0.9 according to κ statistics in an experimental test site in the Chinese city of Dongying. In addition, different comparison methods were evaluated in Chapter III, whereas four out of seven partitioning clustering algorithms were found generally suitable in the experimental change detection setting in the city of Dongying.

Compared to existing studies in literature, the change detection approach in this thesis returned robust results and viable accuracies, which are similar or beyond related work (e.g., Klonus et al. (2012), Huang et al. (2014), Tang et al. (2013), Xiao et al. (2017)). In addition, these studies rely on VHR imagery form a single sensor with similar or identical parameters of the acquisition system (e.g., sensor view angle, etc.), whereas the proposed approach demonstrated its capabilities for heterogeneous input imagery. Most recently, heterogeneous data settings based on VHR image pairs from different sensors were investigated based on different pixel- as well as

object-based approaches by e.g., Jabari et al. (2019), Solano-Correa et al. (2018), Tang and Zhang (2017), or Wang et al. (2015), who achieved similar results and accuracies. In contrast to related studies that employ data with differences in the acquisition system for supervised change detection, the unsupervised approach in this thesis does not require additional a priori or ground truth information and thus, facilitates transferability as well as efficient analysis of large coverage of VHR data.

However, the change detection approach as presented in this thesis has several limitations and additional aspects remain that may be investigated in future research. The approach in this thesis relies on an extensive data setting, which is still quite unique in many areas worldwide, although the availability of VHR data is steadily increasing. Nevertheless, a substitution of data is desirable and will enhance the practical applicability of the change detection methodology. For example, object geometries are generated from an nDSM in this thesis but could alternatively originate from official data sources or volunteered geographic information. Despite, such sources of data remain challenging in terms of availability and quality especially in emerging and developing countries. Furthermore, the transferability to other less dynamic geographical regions is subject to uncertainties, where the distribution of classes plays a major role as shown in Chapter IV and below. To address these uncertainties, other algebraic and arithmetic methods could be considered, in particular with non-linear classification capabilities. In this context, emerging methods of ensemble or deep learning could be taken into consideration for more complex data and change detection settings. Besides the application of urban growth in this thesis, other fields of application such as shrinking cities or damage assessment of buildings are generally feasible based on the proposed approach if a suitable data setting is available. However, the demonstration of its practical applicability is beyond the scope of this thesis and must be examined in future research.

With regard to differences in the acquisition system as well as inevitable inaccuracies of georeferencing, the change detection approach of this thesis offers significant advantages for mitigating these variations and thus, enables proper identification of changes. This is of particular importance with respect to data acquired e.g., by different sensors, with deviating viewing geometries, in different acquisition modes, and/or in different seasons. Nevertheless, hardly any studies exist on this aspect in literature and thus, this thesis contributes significantly to bridge the research gap concerning the usability of heterogeneous VHR data. In this regard, the influence of deviating viewing geometries was examined in Chapter II and showed that experimentally adjusted object geometries improved the change detection result only slightly. Furthermore, monitoring urban areas at medium and small spatial scale will only be feasible by integrating heterogeneous VHR data from all available sources, which in turn will only be possible by appropriate capabilities of handling such heterogeneous data. In addition, the unsupervised implementation offers high potential for automation in order to enable efficient processing of

the increasing amount of VHR imagery. As a methodological extension, different unsupervised clustering algorithms were evaluated and identified in Chapter III, which depicts a first step towards increasing flexibility and robustness of the proposed change detection approach.

2) Evaluation of transferability with regard to the distribution of classes

The transferability of the change detection approach in this thesis was evaluated in Chapter IV with particular focus on the nature and effects of class distribution and potentially occurring class imbalance. For this purpose, a diagnostic framework was developed and consequently applied for the detection of changes in two exemplary study sites with different characteristics of the built environment as well as divergent temporal evolution: the city of Dongying, China representing a highly dynamic urban situation and the city of Munich, Germany representing an urban morphology of comparatively low dynamics. The methodology for change detection directly builds upon Chapter II, whereas several suitable clustering algorithms were tested and investigated further according to the findings in Chapter III. As expected, results showed that situations of imbalanced class distribution generally provide less reliable identification of changes compared to balanced or close to balanced situations. In detail, genetic k-means was found more robust compared to the original k-means algorithm, and the self-organizing map (SOM) clustering was identified as the most prone technique to unevenly distributed classes. In addition, the results of this thesis revealed the sensitivity of various object-based features towards the complexity of the built-up structure as well as the variety of building types. The precision of change detection was evaluated using three different measures of classification accuracy: Cohen's κ coefficient, True Skill Statistic (TSS), and Precision-Recall (PR) curves.

Existing studies in literature related to classification of remote sensing images with particular consideration of class imbalance (e.g., Bruzzone and Serpico (1997), García et al. (2011), Kellenberger et al. (2018), Kubat et al. (1998), Williams et al. (2009)) present case studies in different fields of application aiming at supervised methodological solutions. In contrast, the results presented in this thesis target the description and systematic sensitivity assessment of the distribution and possibly imbalanced classes in an unsupervised change detection context. In accordance with the findings of Xu and Wunsch (2010), algorithms like genetic k-means that account for non-linearity and are able to detect clusters of arbitrary shape provide more robust results with regard to class imbalance in general. As demonstrated in this thesis, the original k-means algorithm is more sensitive to unfavorable data distributions as well as outliers, which is also the case for SOM clustering. In addition, k-means seeks for clusters of similar size (Jain 2010), which deteriorates the change detection results particularly in case of imbalanced classes.

This thesis demonstrated that none of the investigated clustering algorithms provides results with sufficient accuracy in situations with strongly imbalanced distribution of classes (i.e., majority of objects changed or unchanged) based on the proposed change detection approach. Although there are different over- and undersampling methods and several supervised approaches for the classification of imbalanced data (He and Garcia 2009; López et al. 2013), further investigation and development of methods for unsupervised change detection based on VHR remote sensing with respect of class imbalance is highly required. This is of great importance, particularly in terms of transferability of the approach and application of change detection in different geographic areas, different temporal and spatial scales, as well as in different types of application. Therefore, suitable algorithms must be investigated in future research, whereas distance-based approaches for novelty detection appear promising according to Pimentel et al. (2014).

It is evident that is the transferability is a crucial aspect of any change detection or image classification approach. In particular, robustness towards the distribution of the classes of interest is of special importance since it is unknown a priori and varies greatly with the type of application, geographical setting, temporal scale, as well as the spatial extent of change detection. In general, the framework for diagnostic analysis of sensitivity towards class imbalance contributes to research for the evaluation of model transferability in any two-class classification problem. Although the development of a solution for imbalanced class distributions in unsupervised change detection is beyond the scope of this thesis, it contributes as the first study in literature to the investigation of class imbalance based on remote sensing data from this perspective and provides a consistent framework for its diagnostic analysis. This framework can be utilized in the development and application of present and future methods for change detection and image classification in order to assess and improve robustness and transferability of the methodology.

3) Exemplary application of the assessment of the ghost city phenomenon

A suitable conceptual framework for the assessment of the presence or absence of the ghost city phenomenon based on the change detection methodology using optical VHR remote sensing data was proposed in Chapter V of this thesis. This framework approaches undercapacity with respect to the residential population as one of the key characteristics of a ghost city and relates the physical urban growth to the temporal evolution of population. The methodology for change detection (Chapter II) was transferred from the experimental test site to the complete city extent of Dongying using different clustering algorithms (Chapter III) and was complemented by supervised building types classification as well as estimation of the number of floors to establish a 4d functional city model. The 4d functional city model based on VHR remote sensing data was used for multi-temporal

population capacity estimation, which was subsequently opposed to multi-temporal census counts for the assessment of the presence or absence of the ghost city phenomenon. This citywide analysis revealed a massive mismatch of estimated population capacity against population counts over time, which depicts a high likelihood for the emergence and presence of the ghost city phenomenon within the urban area of Dongying. In addition, a spatially more detailed evaluation was conducted in terms of an index comparing the dynamics of residential developments and population numbers, which identified regions within the urban area that are most likely to suffer from the ghost city phenomenon.

To date, similar, but few approaches based on HR remote sensing data have been proposed in literature for the assessment of the ghost city phenomenon (Ma et al. 2018; Zheng et al. 2017a; Zheng et al. 2017b), while this thesis is based on VHR data. The 4d functional city model with its components of the multi-temporal building model, classification of building types, and estimation of the number of floors provides similar accuracies compared to literature (e.g., Belgiu et al. (2014), Geiß et al. (2015a), Huang et al. (2017), Lefebvre and Corpetti (2017), Wurm et al. (2016)). Regarding the estimation of population capacity using a statistical approach according to residential living space, similar work can be found in literature (e.g., Tomás et al. (2016), Wang et al. (2016b)), however undercapacity with respect to population as the unique feature of a ghost city does not allow a direct comparison.

However, there are several limitations and uncertainties which could be investigated for more reliable assessment of the ghost city phenomenon in future research. The underlying 4d functional city model could be improved with respect to the identification of changes of buildings as discussed above, as well as concerning the classification of building types and the estimation of the number of floors. Building types associated with residential and non-residential function may be determined by emerging supervised classification methods such as deep learning, alternatively unsupervised techniques could also be explored. The number of floors is modeled in a uniform linear manner for all residential buildings in this thesis, whereas a more detailed distinction of residential building types together with an individually adjusted transformation of height will likely improve results. Consequently, the estimation of population capacity based on the 4d functional city model would benefit from these improvements and extensions. In addition, official census counts suffer from inherent uncertainties as well as specific errors due to floating population from rural migrant workers in China. From a more general perspective, additional aspects on the economic development or human activities should be included for a broader assessment of the ghost city phenomenon in future work. Furthermore, the methodology for the assessment of the ghost city phenomenon must be applied to different emerging and fast growing cities or at the national scale of China if appropriate VHR data is available.

Nevertheless, the assessment of the ghost city phenomenon in this thesis provides a reasonable framework for the multi-temporal confrontation of available living space of residential buildings (i.e., population capacity) and population counts from official census data. This thesis demonstrates its capabilities for the identification of the ghost city phenomenon based on VHR data with reasonable reliability as well as its more detailed spatial characteristics, whereas regions of the urban area could be identified that are most likely affected. The highly detailed 4d functional city model enables an exact estimation of the dynamics of the urban morphology as well as associated residential and non-residential function, which was demonstrated to be well suited for the assessment of the ghost city phenomenon.

2. Main conclusions

In view of its overall aim, this dissertation demonstrates the applicability of multi-temporal VHR remote sensing data for unsupervised change detection and subsequent assessment of the ghost city phenomenon as an exemplary application from the field of urban geography.

The change detection methodology presented in Chapter II of this thesis has demonstrated its flexibility and thus its ability for monitoring of urban growth in an urban environment based on heterogeneous VHR images from different sensors. Based on its high potential for automation due to unsupervised implementation, it offers suitable capabilities for the detection of changes of the built environment over large areas. This is also of particular relevance in view of the future increasing availability of VHR remote sensing data. In this context, the results regarding the transferability of the approach in Chapter IV can be used to continuously improve the robustness of the change detection methodology. Potential improvement of the robustness was already demonstrated in Chapter III, where different partitioning clustering algorithms were evaluated and distinct characteristics could be identified. In particular, the capabilities of the change detection methodology concerning class imbalance can be investigated and significantly extended based on the diagnostic framework proposed in this thesis and new techniques can be elaborated on its basis. The exemplary application of the ghost city phenomenon in Chapter V is an important aspect of urban growth in the context of sustainable urban development and goes beyond the analysis of urban growth (Si Salah et al. 2019; Zhu et al. 2019). Therefore, the results of this thesis contribute to a more detailed understanding of challenges related to urban growth in China, especially in fast growing small and medium sized Chinese cities like the city of Dongying. The ghost city assessment demonstrates the added value of multi-temporal VHR remote sensing imagery in combination with

ancillary data of census counts for identification and analysis of this aspect of urban growth in China. Previously, this phenomenon has hardly been analyzed and quantified spatially (Shepard 2015), although its increasing importance has been emphasized in several studies (e.g., He et al. (2016), Sorace and Hurst (2016), Woodworth and Wallace (2017)) with particular emphasis on sustainable urbanization as required by the Sustainable Development Goals (i.e., particularly Goal 11) by the United Nations. In addition, these developments may trigger an emerging real estate bubble and debt crisis of local governments, which are themselves increasingly dependent on revenues and taxes from real estate (Jiang et al. 2017). Therefore, sufficient data for persistent monitoring of cities are urgently required for reliable conclusions about trends in urban growth of China in order to identify ghost city prone urban areas and enable appropriate countermeasures to be taken. Moreover, also other fields of application in the context of changing urban environments could be enabled and explored based on the methodology and findings of this thesis in future work.

3. Outlook and future research

This thesis provides a flexible and robust methodology for unsupervised identification of changed buildings in VHR remote sensing images from different sensors and subsequently documents its applicability in an exemplary application in the context of the ghost city phenomenon in the city of Dongying, China. Besides the findings within this thesis, there are several extensions and perspectives for future research that are beyond the scope of this dissertation.

In the context of methodological development, future research must focus on further advancement of robust and flexible change detection approaches that are able to efficiently handle the increasing amount of VHR remote sensing imagery. Special attention must continuously be paid to corresponding capabilities with respect to heterogeneous VHR data with differences in the acquisition system, since the analysis of large spatial coverages will only be feasible by integrating VHR data from all available sources. In addition, also emerging sources of VHR remote sensing data (e.g., nano- and microsatellites like Planet SkySat) should be considered in order to enlarge the data basis for monitoring of the Earth's surface. In this regard, the object-based approach for unsupervised change detection in this thesis has proven its applicability, however numerous other algorithms besides the investigated clustering techniques exist that might improve results. Different comparison methods also enable the identification of the type of change, which would allow a more detailed description of urban dynamics. In this context, the benefit of additional object-based features as well as corresponding preprocessing

methods could be analyzed. Another possible extension is the inclusion of pixel-based metrics for more detailed characterization within objects for change detection. In terms of transferability, this thesis demonstrated that unsupervised methods for imbalanced distribution of classes are urgently required, since such situations are common in change detection within urban environments. In order to address imbalanced distribution of classes, the diagnostic framework in this thesis can be used for development and evaluation of new techniques as well as improvement of the proposed change detection approach. The exemplary application of the assessment of the ghost city phenomenon presents the successful transfer of the change detection methodology to greater spatial coverage and combines changes of land cover (i.e., individual buildings) with ancillary data from census for a new field of application based on VHR data. Although the 4d functional city model provides good accuracy, it might be further enhanced by an improved methodology for change detection as described above, or by means of emerging supervised as well as unsupervised classification techniques for identification of building types and estimation of the number of floors. In addition, a more detailed functional characterization, possibly with different types of residential buildings, will likely improve results. However, a crucial aspect is the accuracy and reliability of ancillary data, which cannot be influenced a priori and must be mitigated as best as possible. Moreover, census data with highest spatial detail available is desirable and must be collected and included for spatially detailed assessment of the ghost city phenomenon. Nevertheless, continuous monitoring of urban areas is highly required for a final decision on the persistent presence of the ghost city phenomenon and its future development.

Although this thesis demonstrates a methodology for unsupervised change detection in the application context of urban growth, other fields of application such as shrinking cities or damage assessment of buildings caused by natural hazards are generally feasible, provided that a suitable data setting is available. Other specific applications related to urban growth, like the dynamics of informal settlements, may be realized and accompanied by appropriate ancillary data building upon the findings of this thesis. Beyond, the transfer to related fields of application, for example morphologic transformations or functional changes of urban areas, could build upon the 4d functional city model as proposed in this thesis. Adjusted or new concepts will be especially necessary for the multi-temporal assessment of other components of urban environments, like the analysis of urban green spaces dynamics or changes with respect to the urban heat island effect. However in a general context, the variety of possible aspects related to urban dynamics must be investigated and analyzed in detail in future research in order to fully capture and understand the evolution of urban systems and ensure sustainable growth of cities.

Against the background of global change and the continuously increasing world population, urban areas and cities will remain hot spots of transformation and ongoing changes across the globe. Remote sensing has proven an appropriate tool for the documentation of these tremendous transformations at unprecedented spatial resolution and current research is able to provide objective and scientific profound knowledge of the processes involved as demonstrated in this thesis. This information must be accessible and used by political decision makers and urban planning in order to pursue and implement the envisaged goal of efficient, resilient, and sustainable development of cities.

References

Akbani, R., Kwek, S., & Japkowicz, N. (2004). Applying Support Vector Machines to Imbalanced Datasets. In J.-F. Boulicaut, F. Esposito, F. Giannotti, & D. Pedreschi (Eds.), *Machine Learning: ECML 2004: 15th European Conference on Machine Learning, Pisa, Italy, September 20-24, 2004. Proceedings* (pp. 39-50). Berlin, Heidelberg: Springer Berlin Heidelberg

Al-Khudhairy, D.H.A., Caravaggi, I., & Glada, S. (2005). Structural damage assessments from Ikonos data using change detection, object-oriented segmentation, and classification techniques. *Photogrammetric Engineering & Remote Sensing, 71*, 825-837

Alavipanah, S., Schreyer, J., Haase, D., Lakes, T., & Qureshi, S. (2018). The effect of multi-dimensional indicators on urban thermal conditions. *Journal of Cleaner Production, 177*, 115-123

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology, 43*, 1223-1232

Aravena Pelizari, P., Spröhnle, K., Geiß, C., Schoepfer, E., Plank, S., & Taubenböck, H. (2018). Multi-sensor feature fusion for very high spatial resolution built-up area extraction in temporary settlements. *Remote Sensing of Environment, 209*, 793-807

Arino, O., Gross, D., Ranera, F., Leroy, M., Bicheron, P., Brockman, C., Defourny, P., Vancutsem, C., Achard, F., Durieux, L., Bourg, L., Latham, J., Gregorio, A.D., Witt, R., Herold, M., Sambale, J., Plummer, S., & Weber, J.L. (2007). GlobCover: ESA service for global land cover from MERIS. In, *2007 IEEE International Geoscience and Remote Sensing Symposium* (pp. 2412-2415)

Balk, D., Pozzi, F., Yetman, G., Deichmann, U., & Nelsonc, A. (2005). The distribution of people and the dimension of place: Methodologies to improve the global estimation of urban extents. In, *Urban Remote Sensing Conference*

Ban, Y., Webber, L., Gamba, P., & Paganini, M. (2017). EO4Urban: Sentinel-1A SAR and Sentinel-2A MSI data for global urban services. In, *2017 Joint Urban Remote Sensing Event (JURSE)* (pp. 1-4)

Baosheng, W., Zhaoyin, W., & Changzhi, L. (2004). Yellow River Basin management and current issues. *Journal of Geographical Sciences, 14*, 29-37

Bartholomé, E., & Belward, A.S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing, 26*, 1959-1977

Bechtel, B., Alexander, P.J., Böhner, J., Ching, J., Conrad, O., Feddema, J., Mills, G., See, L., & Stewart, I. (2015). Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities. *ISPRS International Journal of Geo-Information, 4*, 199

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31

Belgiu, M., Tomljenovic, I., Lampoltshammer, T., Blaschke, T., & Höfle, B. (2014). Ontology-Based Classification of Building Types Detected from Airborne Laser Scanning Data. *Remote Sensing, 6*, 1347-1366

Bellens, R., Gautama, S., Martinez-Fonte, L., Philips, W., Chan, J.C., & Canters, F. (2008). Improved Classification of VHR Images of Urban Areas Using Directional Morphological Profiles. *IEEE Transactions on Geoscience and Remote Sensing, 46*, 2803-2813

Berkhin, P. (2006). A survey of clustering data mining techniques. In J. Kogan, C. Nicholas, & M. Teboulle (Eds.), *Grouping Multidimensional Data* (pp. 25-71): Springer Berlin Heidelberg

Bhagat, V.S. (2012). Use of Remote Sensing Techniques for Robust Digital Change Detection of Land: A Review. *Recent Patents on Space Technology, 2*, 123-144

Biljecki, F., Arroyo Ohori, K., Ledoux, H., Peters, R., & Stoter, J. (2016). Population Estimation Using a 3D City Model: A Multi-Scale Country-Wide Study in the Netherlands. *PLOS ONE, 11*, e0156808

Birkmann, J., Cardona, O.D., Carreño, M.L., Barbat, A.H., Pelling, M., Schneiderbauer, S., Kienberger, S., Keiler, M., Alexander, D., Zeil, P., & Welle, T. (2013). Framing vulnerability, risk and societal responses: the MOVE framework. *Natural Hazards, 67*, 193-211

Blaschke, T. (2005). Towards a framework for change detection based on image objects. *Göttinger Geographische Abhandlungen, 113*, 1-9

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 2-16

Bontemps, S., Bogaert, P., Titeux, N., & Defourny, P. (2008). An object-based change detection method accounting for temporal dependences in time series with medium to coarse spatial resolution. *Remote Sensing of Environment, 112*, 3181-3191

Bontemps, S., Defourny, P., van Bogaert, E., Arino, O., Kalogirou, V., & Perez, J.R. (2010). *GLOBCOVER 2009: Products Description and Validation Report*. Paris, France: European Space Agency (ESA)

Bouziani, M., Goïta, K., & He, D.-C. (2010). Automatic change detection of buildings in urban environment from very high spatial resolution images using existing geodatabase and prior knowledge. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*, 143-153

Bovolo, F., & Bruzzone, L. (2007a). A Split-Based Approach to Unsupervised Change Detection in Large-Size Multitemporal Images: Application to Tsunami-Damage Assessment. *IEEE Transactions on Geoscience and Remote Sensing, 45*, 1658-1670

Bovolo, F., & Bruzzone, L. (2007b). A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain. *IEEE Transactions on Geoscience and Remote Sensing, 45*, 218-236

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5-32

Brunner, D., Lemoine, G., & Bruzzone, L. (2010). Earthquake Damage Assessment of Buildings Using VHR Optical and SAR Imagery. *IEEE Transactions on Geoscience and Remote Sensing, 48*, 2403-2420

Bruzzone, L., & Bovolo, F. (2013). A novel framework for the design of change-detection-systems for very-high-resolution remote sensing images. *Proceedings of the IEEE, 101*, 609-630

Bruzzone, L., & Serpico, S.B. (1997). Classification of imbalanced remote-sensing data by neural networks. *Pattern Recognition Letters, 18*, 1323-1328

Buiten, H.J., & Clevers, J.G.P.W. (1994). *Land Observation by Remote Sensing*. London, United Kingdom: Taylor & Francis

Bullock, E.L., Woodcock, C.E., & Holden, C.E. (2019). Improved change monitoring using an ensemble of time series algorithms. *Remote Sensing of Environment, in press*

Butler, D. (2014). Many eyes on Earth - Swarms of small satellites set to deliver close to real-time imagery of swathes of the planet. *Nature News, 505*, 143-144

Cabrera-Barona, P., Wei, C., & Hagenlocher, M. (2016). Multiscale evaluation of an urban deprivation index: Implications for quality of life and healthcare accessibility planning. *Applied Geography, 70*, 1-10

Cao, G., Wang, B., Xavier, H.-C., Yang, D., & Southworth, J. (2017). A new difference image creation method based on deep neural networks for change detection in remote-sensing images. *International Journal of Remote Sensing, 38*, 7161-7175

Carleer, A.P., & Wolff, E. (2006). Urban land cover multi-level region-based classification of VHR data by selecting relevant features. *International Journal of Remote Sensing, 27*, 1035-1051

Celik, T. (2009). Unsupervised change detection in satellite images using principal component analysis and k-means clustering. *IEEE Geoscience and Remote Sensing Letters, 6*, 772-776

Chan, K.W. (2012). Migration and development in China: trends, geography and current issues. *Migration and Development, 1*, 187-205

Chan, K.W., & Buckingham, W. (2008). Is China Abolishing the Hukou System? *The China Quarterly, 195*, 582-606

Chawla, N.V., Bowyer, K.W., Hall, L.O., & Kegelmeyer, W.P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal Of Artificial Intelligence Research, 16*, 321-357

Chawla, N.V., Japkowicz, N., & Kotcz, A. (2004). Editorial: special issue on learning from imbalanced data sets. *ACM SIGKDD Explorations Newsletter - Special issue on learning from imbalanced datasets, 6*, 1-6

Chen, G., Hay, G.J., Carvalho, L.M.T., & Wulder, M.A. (2012). Object-based change detection. *International Journal of Remote Sensing, 33*, 4434-4457

Chen, J., Chang, K.-t., Karacsonyi, D., & Zhang, X. (2014). Comparing urban land expansion and its driving factors in Shenzhen and Dongguan, China. *Habitat International, 43*, 61-71

Chen, Q., & Chen, Y. (2016). Multi-Feature Object-Based Change Detection Using Self-Adaptive Weight Change Vector Analysis. *Remote Sensing, 8*, 549-568

Chen, Z., & Hutchinson, T.C. (2007). Urban damage estimation using statistical processing of satellite images. *Journal of Computing in Civil Engineering, 21*, 187-199

Chi, G., Liu, Y., & Wu, H. (2015). *Ghost Cities Analysis Based on Positioning Data in China*. arXiv preprint arXiv:1510.08505

Clausi, D.A. (2002). An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing, 28*, 45-62

Cohen, B. (2006). Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society, 28*, 63-80

Comitee of the Dongying urban and rural planning bureau (2011). *Dongying urban and rural planning*. China Publishing House

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35-46

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., & Lambin, E. (2004). Review article - Digital change detection methods in ecosystem monitoring: A review. *International Journal of Remote Sensing, 25*, 1565-1596

Crutzen, P.J. (2002). Geology of mankind. *Nature, 415*, 23

Cui, B., Yang, Q., Yang, Z., & Zhang, K. (2009). Evaluating the ecological performance of wetland restoration in the Yellow River Delta, China. *Ecological Engineering, 35*, 1090-1103

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. In, *Proceedings of the 23rd international conference on Machine learning* (pp. 233-240). Pittsburgh, Pennsylvania, USA: ACM

Dell'Acqua, F., & Gamba, P. (2012). Remote Sensing and Earthquake Damage Assessment: Experiences, Limits, and Perspectives. *Proceedings of the IEEE, 100*, 2876-2890

Desclée, B., Bogaert, P., & Defourny, P. (2006). Forest change detection by statistical object-based method. *Remote Sensing of Environment, 102*, 1-11

Di, S., Li, Z.-L., Tang, R., Pan, X., Liu, H., & Niu, Y. (2019). Urban green space classification and water consumption analysis with remote-sensing technology: a case study in Beijing, China. *International Journal of Remote Sensing, 40*, 1909-1929

Ding, K., Huo, C., Xu, Y., Zhong, Z., & Pan, C. (2015). Sparse hierarchical clustering for VHR image change detection. *IEEE Geoscience and Remote Sensing Letters, 12*, 577-581

Dobson, J., A. Bright, E., R. Coleman, P., C. Durfee, R., & A. Worley, B. (2000). LandScan: A Global Population Database for Estimating Populations at Risk. *Photogrammetric Engineering and Remote Sensing, 66*, 849-857

Doncaster, B., Shulman, J., Bradford, J., & Olds, J. (2016). SpaceWorks' 2016 Nano/Microsatellite Market Forcast

Doxani, G., Karantzalos, K., & Strati, M.T. (2012). Monitoring urban changes based on scale-space filtering and object-oriented classification. *International Journal of Applied Earth Observation and Geoinformation, 15*, 38-48

Doxsey-Whitfield, E., MacManus, K., Adamo, S.B., Pistolesi, L., Squires, J., Borkovska, O., & Baptista, S.R. (2015). Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4. *Papers in Applied Geography, 1*, 226-234

Dueker, K.J., & Horton, F.E. (1972). Urban-change detection systems: Remote-sensing inputs. *Photogrammetria, 28*, 89-106

Durieux, L., Lagabrielle, E., & Nelson, A. (2008). A method for monitoring building construction in urban sprawl areas using object-based analysis of Spot 5 images and existing GIS data. *ISPRS Journal of Photogrammetry and Remote Sensing, 63*, 399-408

Eckert, S., Jelinek, R., Zeug, G., & Krausmann, E. (2012). Remote sensing-based assessment of tsunami vulnerability and risk in Alexandria, Egypt. *Applied Geography, 32*, 714-723

Editorial Committee of Dongying Statistical Yearbook (2013). *Dongying statistical yearbook* Beijing: China Statistics Press

Elkan, C. (2001). The foundations of cost-sensitive learning. In, *Proceedings of the 17th international joint conference on Artificial intelligence - Volume 2* (pp. 973-978). Seattle, WA, USA: Morgan Kaufmann Publishers Inc.

Ertekin, S., Huang, J., Bottou, L., & Giles, L. (2007). Learning on the border: active learning in imbalanced data classification. In, *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management* (pp. 127-136). Lisbon, Portugal: ACM

Esch, T., Bachofer, F., Heldens, W., Hirner, A., Marconcini, M., Palacios-Lopez, D., Roth, A., Üreyen, S., Zeidler, J., Dech, S., & Gorelick, N. (2018). Where We Live—A Summary of the Achievements and Planned Evolution of the Global Urban Footprint. *Remote Sensing, 10*, 895-913

Esch, T., Marconcini, M., Felbier, A., Roth, A., Heldens, W., Huber, M., Schwinger, M., Taubenböck, H., Müller, A., & Dech, S. (2013). Urban Footprint Processor - Fully Automated Processing Chain Generating Settlement Masks From Global Data of the TanDEM-X Mission. *IEEE Geoscience and Remote Sensing Letters, 10*, 1617-1621

Falco, N., Mura, M.D., Bovolo, F., Benediktsson, J.A., & Bruzzone, L. (2013). Change Detection in VHR Images Based on Morphological Attribute Profiles. *IEEE Geoscience and Remote Sensing Letters, 10*, 636-640

Fang Ranking (2015). Top 12 "ghost cities" of China in 2015. In. http://news.anshan.fang.com/2015-12-08/18506456_all.htm

Ferré, L. (1995). Selection of components in principal component analysis: A comparison of methods. *Computational Statistics & Data Analysis, 19*, 669-682

Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*, 185-201

Fraley, C., & Raftery, A.E. (2002). Model-Based Clustering, Discriminant Analysis, and Density Estimation. *Journal of the American Statistical Association, 97*, 611-631

Gan, G., Ma, C., & Wu, J. (2007). *Data clustering - Theory, algorithms, and applications*. Philadelphia, Alexandria: SIAM, ASA

Gar-on Yeh, A., & Wu, F. (1999). The transformation of the urban planning system in China from a centrally-planned to transitional economy. *Progress in Planning, 51*, 167-252

García, V., Sánchez, J.S., & Mollineda, R.A. (2011). Classification of High Dimensional and Imbalanced Hyperspectral Imagery Data. In J. Vitrià, J.M. Sanches, & M. Hernández (Eds.), *Pattern Recognition and Image Analysis: 5th Iberian Conference, IbPRIA 2011, Las Palmas de Gran Canaria, Spain, June 8-10, 2011. Proceedings* (pp. 644-651). Berlin, Heidelberg: Springer Berlin Heidelberg

Geiß, C., Aravena Pelizari, P., Marconcini, M., Sengara, W., Edwards, M., Lakes, T., & Taubenböck, H. (2015a). Estimation of seismic building structural types using multi-sensor remote sensing and machine learning techniques. *ISPRS Journal of Photogrammetry and Remote Sensing, 104*, 175-188

Geiß, C., Taubenböck, H., Wurm, M., Esch, T., Nast, M., Schillings, C., & Blaschke, T. (2011). Remote Sensing-Based Characterization of Settlement Structures for Assessing Local Potential of District Heat. *Remote Sensing, 3*, 1447-1471

Geiß, C., Wurm, M., Breunig, M., Felbier, A., & Taubenböck, H. (2015b). Normalization of TanDEM-X DSM data in urban environments with morphological filters. *IEEE Transactions on Geoscience and Remote Sensing, 53*, 4348-4362

Ghosh, A., Mishra, N.S., & Ghosh, S. (2011). Fuzzy clustering algorithms for unsupervised change detection in remote sensing images. *Information Sciences, 181*, 699-715

Gong, J., Sui, H., Sun, K., Ma, G., & Liu, J. (2008). Object-level change detection based on full-scale image segmentation and its application to Wenchuan Earthquake. *Science in China Series E: Technological Sciences, 51*, 110-122

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., & Briggs, J.M. (2008). Global Change and the Ecology of Cities. *Science, 319*, 756-760

Gruebner, O., Sachs, J., Nockert, A., Frings, M., Khan, M.M.H., Lakes, T., & Hostert, P. (2014). Mapping the Slums of Dhaka from 2006 to 2010. *Dataset Papers in Science, 2014*, 7

Gueguen, L., & Hamid, R. (2016). Toward a Generalizable Image Representation for Large-Scale Change Detection: Application to Generic Damage Analysis. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 3378-3387

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On Clustering Validation Techniques. *Journal of Intelligent Information Systems, 17*, 107-145

Hall, O., & Hay, G.J. (2003). A Multiscale Object-Specific Approach to Digital Change Detection. *International Journal of Applied Earth Observation and Geoinformation, 4*, 311-327

Handayani, H.H., Murayama, Y., Ranagalage, M., Liu, F., & Dissanayake, D. (2018). Geospatial Analysis of Horizontal and Vertical Urban Expansion Using Multi-Spatial Resolution Data: A Case Study of Surabaya, Indonesia. *Remote Sensing, 10*, 1599-1623

Haralick, R.M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-3*, 610-621

Hartigan, J.A., & Wong, M.A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistics Society. Series C (Applied Statistics), 28*, 100-108

He, G., Mol, A.P.J., & Lu, Y. (2016). Wasted cities in urbanizing China. *Environmental Development, 18*, 2-13

He, H., & Garcia, E.A. (2009). Learning from Imbalanced Data. *Knowledge and Data Engineering, IEEE Transactions on, 21*, 1263-1284

Huang, X., Wen, D., Li, J., & Qin, R. (2017). Multi-level monitoring of subtle urban changes for the megacities of China using high-resolution multi-view satellite imagery. *Remote Sensing of Environment, 196*, 56-75

Huang, X., Zhang, L., & Zhu, T. (2014). Building change detection from multitemporal high-resolution remotely sensed images based on a morphological building index. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 105-115

Huang, Y., Yu, B., Zhou, J., Hu, C., Tan, W., Hu, Z., & Wu, J. (2013). Toward automatic estimation of urban green volume using airborne LiDAR data and high resolution Remote Sensing images. *Frontiers of Earth Science, 7*, 43-54

Hussain, M., Chen, D., Cheng, A., Wei, H., & Stanley, D. (2013). Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing, 80*, 91-106

Im, J., Jensen, J.R., & Tullis, J.A. (2008). Object-based change detection using correlation image analysis and image segmentation. *International Journal of Remote Sensing, 29*, 399-423

IPBES (2019). Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science- Policy Platform on Biodiversity and Ecosystem Services. In J.S. E. S. Brondizio, S. Díaz, and H. T. Ngo (Ed.). Bonn, Germany: IPBES Secretariat

IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In R.K. Pachauri, & L.A. Meyer (Eds.) (p. 151). Geneva, Switzerland: Intergovernmental Panel on Climate Change

Jabari, S., Rezaee, M., Fathollahi, F., & Zhang, Y. (2019). Multispectral change detection using multivariate Kullback-Leibler distance. *ISPRS Journal of Photogrammetry and Remote Sensing, 147*, 163-177

Jackson, D.A. (1993). Stopping rules in principal component analysis: A comparison of heuristical and statistical approaches. *Ecology, 74*, 2204-2214

Jain, A.K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters, 31*, 651-666

Jain, A.K., Murty, M.N., & Flynn, P.J. (1999). Data clustering: A review. *ACM Computing Surveys, 31*, 264-323

Japkowicz, N. (2001). Supervised Versus Unsupervised Binary-Learning by Feedforward Neural Networks. *Machine Learning, 42*, 97-122

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis, 6*, 429-449

Jeni, A., Cohn, J.F., & Torre, F.D.L. (2013). Facing Imbalanced Data Recommendations for the Use of Performance Metrics. In, *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (pp. 245-251): IEEE Computer Society

Jenks, G.F., & Caspall, F.C. (1971). Error on choropletic maps: Definition, measurement, reduction. *Annals of the Association of American Geographers, 61*, 217-244

Jiang, Q.o., Deng, X., Zhan, J., & Yan, H. (2011). Impacts of economic development on ecosystem risk in the Yellow River Delta. *Procedia Environmental Sciences, 5*, 208-218

Jiang, Y., Mohabir, N., Ma, R., & Zhu, P. (2017). Sorting through Neoliberal Variations of Ghost Cities in China. *Land Use Policy, 69*, 445-453

Jin, X., Long, Y., Sun, W., Lu, Y., Yang, X., & Tang, J. (2017). Evaluating cities' vitality and identifying ghost cities in China with emerging geographical data. *Cities, 63*, 98-109

Kantardzic, M. (2011). *Data mining: Concepts, models, methods, and algorithms*. (2nd ed.). Hoboken, New Jersey: John Wiley & Sons

Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab - An S4 Package for Kernel Methods in R. *2004, 11*, 20

Kellenberger, B., Marcos, D., & Tuia, D. (2018). Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning. *Remote Sensing of Environment, 216*, 139-153

Kent, C.W., Grimmond, S., Gatey, D., & Hirano, K. (2019). Urban morphology parameters from global digital elevation models: Implications for aerodynamic roughness and for wind-speed estimation. *Remote Sensing of Environment, 221*, 316-339

Kit, O., & Lüdeke, M. (2013). Automated detection of slum area change in Hyderabad, India using multitemporal satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 83*, 130-137

Klaric, M.N., Claywell, B.C., Scott, G.J., Hudson, N.J., Sjahputera, O., Li, Y., Barratt, S.T., Keller, J.M., & Davis, C.H. (2013). GeoCDX: An Automated Change Detection and Exploitation System for High-Resolution Satellite Imagery. *IEEE Transactions on Geoscience and Remote Sensing, 51*, 2067-2086

Klonus, S., Tomowski, D., Ehlers, M., Reinartz, P., & Michel, U. (2012). Combined edge segment texture analysis for the detection of damaged buildings in crisis areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*, 1118-1128

Klotz, M., Kemper, T., Geiß, C., Esch, T., & Taubenböck, H. (2016). How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sensing of Environment, 178*, 191-212

Knoth, C., & Pebesma, E. (2017). Detecting dwelling destruction in Darfur through object-based change analysis of very high-resolution imagery. *International Journal of Remote Sensing, 38*, 273-295

Kopsiaftis, G., & Karantzalos, K. (2015). Vehicle detection and traffic density monitoring from very high resolution satellite video data. In, *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1881-1884)

Kranz, O., Lang, S., & Schoepfer, E. (2017). 2.5D change detection from satellite imagery to monitor small-scale mining activities in the Democratic Republic of the Congo. *International Journal of Applied Earth Observation and Geoinformation, 61*, 81-91

Kubat, M., Holte, R.C., & Matwin, S. (1998). Machine Learning for the Detection of Oil Spills in Satellite Radar Images. *Machine Learning, 30*, 195-215

Kuenzer, C., Ottinger, M., Liu, G., Sun, B., Baumhauer, R., & Dech, S. (2014). Earth observation-based coastal zone monitoring of the Yellow River Delta: Dynamics in China's second largest oil producing region over four decades. *Applied Geography, 55*, 92-107

Kukar, M., & Kononenko, I. (1998). Cost-Sensitive Learning with Neural Networks. In, *13th European Conference on Artificial Intelligence* (pp. 445-449)

Lee, H.-j., & Cho, S. (2006). The novelty detection approach for different degrees of class imbalance. In, *Proceedings of the 13th international conference on Neural Information Processing - Volume Part II* (pp. 21-30). Hong Kong, China: Springer-Verlag

Lefebvre, A., & Corpetti, T. (2017). Monitoring the Morphological Transformation of Beijing Old City Using Remote Sensing Texture Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*, 539-548

Lefebvre, A., Corpetti, T., & Hubert-Moy, L. (2008). Object-Oriented Approach and Texture Analysis for Change Detection in Very High Resolution Images. In, *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium* (pp. IV - 663-IV - 666)

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., & Taubenböck, H. (2017a). Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment. *International Journal of Applied Earth Observation and Geoinformation, 54*, 15-27

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., & Taubenböck, H. (2017b-a). Evaluation of clustering algorithms for unsupervised change detection in VHR remote sensing imagery. In, *2017 Joint Urban Remote Sensing Event (JURSE)*. Dubai, UAE

Leichtle, T., Geiß, C., Wurm, M., Lakes, T., & Taubenböck, H. (2017b-b). Evaluation of clustering algorithms for unsupervised change detection in VHR remote sensing imagery. In, *2017 Joint Urban Remote Sensing Event (JURSE)* (pp. 1-4)

Lewis, S.L., & Maslin, M.A. (2015). Defining the Anthropocene. *Nature, 519*, 171

Li, M., Stein, A., Bijker, W., & Zhan, Q. (2016). Urban land use extraction from Very High Resolution remote sensing imagery using a Bayesian network. *ISPRS Journal of Photogrammetry and Remote Sensing, 122*, 192-205

Li, P., Xu, H., & Guo, J. (2010). Urban building damage detection from very high resolution imagery using OCSVM and spatial features. *International Journal of Remote Sensing, 31*, 3393-3409

Li, Q., Mou, L., Xu, Q., Zhang, Y., & Zhu, X.X. (2019). R³-Net: A Deep Network for Multioriented Vehicle Detection in Aerial Images and Videos. *IEEE Transactions on Geoscience and Remote Sensing, 57*, 5028-5042

Li, X., & Gong, P. (2016). An “exclusion-inclusion” framework for extracting human settlements in rapidly developing regions of China from Landsat images. *Remote Sensing of Environment, 186*, 286-296

Li, X., Gong, P., & Liang, L. (2015). A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data. *Remote Sensing of Environment, 166*, 78-90

Li, X., Zhao, S., Yang, H., Cong, D., & Zhang, Z. (2017a). A Bi-Band Binary Mask Based Land-Use Change Detection Using Landsat 8 OLI Imagery. *Sustainability, 9*, 479-494

Li, Z., Shi, W., Hao, M., & Zhang, H. (2017b). Unsupervised change detection using spectral features and a texture difference measure for VHR remote-sensing images. *International Journal of Remote Sensing, 38*, 7302-7315

Li, Z., Zhou, C., Yang, X., Chen, X., Meng, F., Lu, C., Pan, T., & Qi, W. (2018). Urban landscape extraction and analysis in the mega-city of China's coastal regions using high-resolution satellite imagery: A case of Shanghai, China. *International Journal of Applied Earth Observation and Geoinformation, 72*, 140-150

Liang, Z. (2001). The Age of Migration in China. *Population and Development Review, 27*, 499-524

Liang, Z., Li, Z., & Ma, Z. (2014). Changing Patterns of the Floating Population in China during 2000-2010. *Population and Development Review, 40*, 695-716

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news, 2*, 18-22

Lin, T., Sun, C., Li, X., Zhao, Q., Zhang, G., Ge, R., Ye, H., Huang, N., & Yin, K. (2016). Spatial pattern of urban functional landscapes along an urban–rural gradient: A case study in Xiamen City, China. *International Journal of Applied Earth Observation and Geoinformation, 46*, 22-30

Listner, C., & Niemeyer, I. (2011). Recent advances in object-based change detection. In, *2011 IEEE International Geoscience and Remote Sensing Symposium* (pp. 110-113)

Liu, G., Zhang, L., Zhang, Q., Musyimi, Z., & Jiang, Q. (2014). Spatio–Temporal Dynamics of Wetland Landscape Patterns Based on Remote Sensing in Yellow River Delta, China. *Wetlands, 34*, 787-801

Liu, H., Yang, M., Chen, J., Hou, J., & Deng, M. (2018a). Line-Constrained Shape Feature for Building Change Detection in VHR Remote Sensing Imagery. *ISPRS International Journal of Geo-Information, 7*, 410-428

Liu, X., Hu, G., Chen, Y., Li, X., Xu, X., Li, S., Pei, F., & Wang, S. (2018b). High-resolution multi-temporal mapping of global urban land using Landsat images based on the Google Earth Engine Platform. *Remote Sensing of Environment, 209*, 227-239

López, V., Fernández, A., García, S., Palade, V., & Herrera, F. (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences, 250*, 113-141

Lu, D., Li, G., & Moran, E. (2014a). Current situation and needs of change detection techniques. *International Journal of Image and Data Fusion, 5*, 13-38

Lu, D., Mausel, P., Brondízio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*, 2365-2401

Lu, Z., Im, J., Rhee, J., & Hodgson, M. (2014b). Building type classification using spatial and landscape attributes derived from LiDAR remote sensing data. *Landscape and Urban Planning, 130*, 134-148

Luo, X., Tong, X., Qian, Z., Pan, H., & Liu, S. (2019). Detecting urban ecological land-cover structure using remotely sensed imagery: A multi-area study focusing on metropolitan inner cities. *International Journal of Applied Earth Observation and Geoinformation, 75*, 106-117

Lv, Z., Liu, T., Shi, C., Benediktsson, J.A., & Du, H. (2019). Novel Land Cover Change Detection Method Based on k-Means Clustering and Adaptive Majority Voting Using Bitemporal Remote Sensing Images. *IEEE Access, 7*, 34425-34437

Ma, T., Zhou, Y., Zhou, C., Haynie, S., Pei, T., & Xu, T. (2015). Night-time light derived estimation of spatio-temporal characteristics of urbanization dynamics using DMSP/OLS satellite data. *Remote Sensing of Environment, 158*, 453-464

Ma, X., Tong, X., Liu, S., Li, C., & Ma, Z. (2018). A Multisource Remotely Sensed Data Oriented Method for "Ghost City" Phenomenon Identification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11*, 2310-2319

Matikainen, L., Hyyppä, J., Ahokas, E., Markelin, L., & Kaartinen, H. (2010). Automatic detection of buildings and changes in buildings for updating of maps. *Remote Sensing, 2*, 1217-1248

Maulik, U., & Bandyopadhyay, S. (2000). Genetic algorithm-based clustering technique. *Pattern Recognition, 33*, 1455-1465

Miller, R.B., & Small, C. (2003). Cities from space: potential applications of remote sensing in urban environmental research and policy. *Environmental Science & Policy, 6*, 129-137

Milligan, G.W., & Cooper, M.C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification, 5*, 181-204

Mou, L., Bruzzone, L., & Zhu, X.X. (2019). Learning Spectral-Spatial-Temporal Features via a Recurrent Convolutional Neural Network for Change Detection in Multispectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing, 57*, 924-935

National Bureau of Statistics of China (2018). *China statistical yearbook*. Beijing: China Statistics Press

Nega, T., Smith, C., Bethune, J., & Fu, W.-H. (2012). An analysis of landscape penetration by road infrastructure and traffic noise. *Computers, Environment and Urban Systems, 36*, 245-256

NetEase Ranking (2015). The urban construction speed exceeds population growth in many cities of China, a large number of "ghost cities" have appeared. In. http://news.163.com/15/1209/01/BABTGM1O00014MTN.html

Niebergall, S., Loew, A., & Mauser, W. (2008). Integrative Assessment of Informal Settlements Using VHR Remote Sensing Data—The Delhi Case Study. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 1*, 193-205

Niemeyer, I., Marpu, P.R., & Nussbaum, S. (2008). Change detection using object features. In T. Blaschke, S. Lang, & G.J. Hay (Eds.), *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications* (pp. 185-201). Berlin, Heidelberg: Springer Berlin Heidelberg

O'Callaghan, C., Boyle, M., & Kitchin, R. (2014). Post-politics, crisis, and Ireland's 'ghost estates'. *Political Geography, 42*, 121-133

Olsen, M.J., Chen, Z., Hutchinson, T., & Kuester, F. (2013). Optical techniques for multiscale damage assessment. *Geomatics, Natural Hazards and Risk, 4*, 49-70

Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich, United Kingdom: Geo Books

Ottinger, M., Clauss, K., & Kuenzer, C. (2016). Aquaculture: Relevance, distribution, impacts and spatial assessments – A review. *Ocean & Coastal Management, 119*, 244-266

Ottinger, M., Kuenzer, C., Liu, G., Wang, S., & Dech, S. (2013). Monitoring land cover dynamics in the Yellow River Delta from 1995 to 2010 based on Landsat 5 TM. *Applied Geography, 44*, 53-68

Park, N.W., & Chi, K.H. (2008). Quantitative assessment of landslide susceptibility using high-resolution remote sensing data and a generalized additive model. *International Journal of Remote Sensing, 29*, 247-264

Patino, J.E., Duque, J.C., Pardo-Pascual, J.E., & Ruiz, L.A. (2014). Using remote sensing to assess the relationship between crime and the urban layout. *Applied Geography, 55*, 48-60

Peres-Neto, P.R., Jackson, D.A., & Somers, K.M. (2005). How many principal components? Stopping rules for determining the number of non-trivial axes revisited. *Computational Statistics & Data Analysis, 49*, 974-997

Pesaresi, M., Gerhardinger, A., & Haag, F. (2007). Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas. *International Journal of Remote Sensing, 28*, 3013-3036

Pesaresi, M., Huadong, G., Blaes, X., Ehrlich, D., Ferri, S., Gueguen, L., Halkia, M., Kauffmann, M., Kemper, T., Lu, L., Marin-Herrera, M.A., Ouzounis, G.K., Scavazzon, M., Soille, P., Syrris, V., & Zanchetta, L. (2013). A Global Human Settlement Layer From Optical HR/VHR RS Data: Concept and First Results. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*, 2102-2131

Pimentel, M.A.F., Clifton, D.A., Clifton, L., & Tarassenko, L. (2014). A review of novelty detection. *Signal Processing, 99*, 215-249

Pittore, M., & Wieland, M. (2013). Toward a rapid probabilistic seismic vulnerability assessment using satellite and ground-based remote sensing. *Natural Hazards, 68*, 115-145

Pontius, R.G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. *International Journal of Remote Sensing, 32*, 4407-4429

Qin, R., Tian, J., & Reinartz, P. (2016). 3D change detection – Approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing, 122*, 41-56

Radhika, S., Tamura, Y., & Matsui, M. (2015). Cyclone damage detection on building structures from pre- and post-satellite images using wavelet based pattern recognition. *Journal of Wind Engineering and Industrial Aerodynamics, 136*, 23-33

Radke, R.J., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing, 14*, 294-307

Ren, X. (2013). *Urban China*. Malden, USA: Polity Press

Richards, J.A. (2012). *Remote Sensing Digital Image Analysis: An Introduction*. Springer Berlin Heidelberg

Richter, R. (1996). A spatially adaptive fast atmospheric correction algorithm. *International Journal of Remote Sensing, 17*, 1201-1214

Ridd, M.K., & Liu, J. (1998). A Comparison of Four Algorithms for Change Detection in an Urban Environment. *Remote Sensing of Environment, 63*, 95-100

Römer, H., Willroth, P., Kaiser, G., Vafeidis, A.T., Ludwig, R., Sterr, H., & Revilla Diez, J. (2012). Potential of remote sensing techniques for tsunami hazard and vulnerability analysis – a case study from Phang-Nga province, Thailand. *Nat. Hazards Earth Syst. Sci., 12*, 2103-2126

Rougier, S., Puissant, A., Stumpf, A., & Lachiche, N. (2016). Comparison of sampling strategies for object-based classification of urban vegetation from Very High Resolution satellite images. *International Journal of Applied Earth Observation and Geoinformation, 51*, 60-73

Ruibo, H., & Linna, W. (2013). Challenges and opportunities facing China's urban development in the new era: A statistical and spatial analysis. *China Perspectives, 2013*, 15-28

Santos, T., Gomes, N., Freire, S., Brito, M.C., Santos, L., & Tenedório, J.A. (2014). Applications of solar mapping in the urban environment. *Applied Geography, 51*, 48-57

Sathyakumar, V., Ramsankaran, R., & Bardhan, R. (2018). Linking remotely sensed Urban Green Space (UGS) distribution patterns and Socio-Economic Status (SES) - A multi-scale probabilistic analysis based in Mumbai, India. *GIScience & Remote Sensing*, 1-25

Schneider, A., Friedl, M.A., & Potere, D. (2009). A new map of global urban extent from MODIS satellite data. *Environmental Research Letters, 4*, 044003

Schneider, A., Friedl, M.A., & Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions'. *Remote Sensing of Environment, 114*, 1733-1746

Schneider, A., & Mertes, C.M. (2014). Expansion and growth in Chinese cities, 1978–2010. *Environmental Research Letters, 9*, 024008

Schölkopf, B., Smola, A., & Müller, K.-R. (1998). Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Computation, 10*, 1299-1319

Scholten, H.J., & Stillwell, J. (2013). *Geographical Information Systems for Urban and Regional Planning*. Springer Netherlands

Schreyer, J., Tigges, J., Lakes, T., & Churkina, G. (2014). Using Airborne LiDAR and QuickBird Data for Modelling Urban Tree Carbon Storage and Its Distribution—A Case Study of Berlin. *Remote Sensing, 6*, 10636-10655

Seto, K.C., Fragkias, M., Güneralp, B., & Reilly, M.K. (2011). A Meta-Analysis of Global Urban Land Expansion. *PLOS ONE, 6*, e23777

Shah-Hosseini, R., Safari, A., & Homayouni, S. (2017). Natural hazard damage detection based on object-level support vector data description of optical and SAR Earth observations *International Journal of Remote Sensing, 38*, 3356-3374

Shangdong Provincial Bureau of Statistics (2001-2016). *Shandong Statistical Yearbook*. Beijing, China: China Statistics Press

Shao, A., Wertz, J.R., & Koltz, E.A. (2017). Quantifying the Cost Reduction Potential for Earth Observation Satellites. In (pp. 199-210). Cham: Springer International Publishing

Shepard, W. (2015). *Ghost Cities of China*. London, United Kingdom: Zed Books

Shi, L., Taubenböck, H., Zhang, Z., Liu, F., & Wurm, M. (2017). Urbanization in China from the end of 1980s until 2010 – spatial dynamics and patterns of growth using EO-data. *International Journal of Digital Earth*, 1-17

Shi, L., Taubenböck, H., Zhang, Z., Liu, F., & Wurm, M. (2019). Urbanization in China from the end of 1980s until 2010 – spatial dynamics and patterns of growth using EO-data. *International Journal of Digital Earth, 12*, 78-94

Si Salah, H., Ait-Aoudia, S., Rezgui, A., & Goldin, S.E. (2019). Change detection in urban areas from remote sensing data: a multidimensional classification scheme. *International Journal of Remote Sensing, 40*, 6635-6679

Simpson, L. (2007). Fixing the Population: From Census to Population Estimate. *Environment and Planning A: Economy and Space, 39*, 1045-1057

Singh, A. (1989). Review article - Digital change detection techniques using remotely-sensed data. *International Journal of Remote Sensing, 10*, 989-1003

Sofina, N., & Ehlers, M. (2016). Building Change Detection Using High Resolution Remotely Sensed Data and GIS. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*, 3430-3438

Solano-Correa, Y.T., Bovolo, F., & Bruzzone, L. (2018). An Approach for Unsupervised Change Detection in Multitemporal VHR Images Acquired by Different Multispectral Sensors. *Remote Sensing, 10*, 533-555

Sorace, C., & Hurst, W. (2016). China's Phantom Urbanisation and the Pathology of Ghost Cities. *Journal of Contemporary Asia, 46*, 304-322

Sowmya, K., John, C.M., & Shrivasthava, N.K. (2015). Urban flood vulnerability zoning of Cochin City, southwest coast of India, using remote sensing and GIS. *Natural Hazards, 75*, 1271-1286

Spence, M., Annez, P.C., & Buckley, R.M. (2009). *Urbanization and Growth*. Washington, DC, USA: World Bank Publications

Spröhnle, K., Fuchs, E., & Pelizari, P.A. (2017). Object-Based Analysis and Fusion of Optical and SAR Satellite Data for Dwelling Detection in Refugee Camps. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10*, 1780-1791

Sripada, R.P., Heiniger, R.W., White, J.G., & Meijer, A.D. (2006). Aerial color infrared photography for determining early in-season nitrogen requirements in corn. *Agronomy Journal, 98*, 968-977

Standard Ranking (2015). "Ghost city" ranking in China. In. http://www.biaozhun007.com/articles/5b3ea9ac91f2fa89.html

Steffen, W., Crutzen, P.J., & McNeill, J.R. (2007). The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature? *Ambio, 36*, 614-621

Stevens, F.R., Gaughan, A.E., Linard, C., & Tatem, A.J. (2015). Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. *PLOS ONE, 10*, e0107042

Stewart, I.D., & Oke, T.R. (2012). Local Climate Zones for Urban Temperature Studies. *Bulletin of the American Meteorological Society, 93*, 1879-1900

Stow, D., Hamada, Y., Coulter, L., & Anguelova, Z. (2008). Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. *Remote Sensing of Environment, 112*, 1051-1061

Sun, Y., Kamel, M.S., Wong, A.K.C., & Wang, Y. (2007). Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognition, 40*, 3358-3378

Tan, M., Li, X., Xie, H., & Lu, C. (2005). Urban land expansion and arable land loss in China—a case study of Beijing–Tianjin–Hebei region. *Land Use Policy, 22*, 187-196

Tanathong, S., Rudahl, K.T., & Goldin, S.E. (2008). Object oriented change detection of buildings after the Indian ocean tsunami disaster. In, *2008 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology* (pp. 65-68)

Tang, Y., Huang, X., & Zhang, L. (2013). Fault-tolerant building change detection from urban high-resolution remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters, 10*, 1060-1064

Tang, Y., & Zhang, L. (2017). Urban Change Analysis with Multi-Sensor Multispectral Imagery. *Remote Sensing, 9*, 252-270

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A., & Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote Sensing of Environment, 117*, 162-176

Taubenböck, H., Goseberg, N., Lämmel, G., Setiadi, N., Schlurmann, T., Nagel, K., Siegert, F., Birkmann, J., Traub, K.-P., Dech, S., Keuck, V., Lehmann, F., Strunz, G., & Klüpfel, H. (2013). Risk reduction at the "Last-Mile": an attempt to turn science into action by the example of Padang, Indonesia. *Natural Hazards, 65*, 915-945

Taubenböck, H., Roth, A., & Dech, S. (2007). Linking structural urban characteristics derived from high resolution satellite data to population distribution In V. Coors, M. Rumor, E. Fendel, & S. Zlatanova (Eds.), *Urban and Regional Data Management* (pp. 35-46): Taylor & Francis Group

Taubenböck, H., & Wiesner, M. (2015). The spatial network of megaregions - Types of connectivity between cities based on settlement patterns derived from EO-data. *Computers, Environment and Urban Systems, 54*, 165-180

Taubenböck, H., Wiesner, M., Felbier, A., Marconcini, M., Esch, T., & Dech, S. (2014). New dimensions of urban landscapes: The spatio-temporal evolution from a polynuclei area to a mega-region based on remote sensing data. *Applied Geography, 47*, 137-153

Taubenböck, H., & Wurm, M. (2015). Ich weiß, dass ich nichts weiß – Bevölkerungsschätzung in der Megacity Mumbai. In H. Taubenböck, M. Wurm, T. Esch, & S. Dech (Eds.), *Globale Urbanisierung: Perspektive aus dem All* (pp. 171-178). Berlin, Heidelberg: Springer Berlin Heidelberg

Taubenböck, H., Wurm, M., Esch, T., Dech, S., Wanka, J., & Wörner, J.D. (2015). *Globale Urbanisierung: Perspektive aus dem All*. Springer Berlin Heidelberg

Taubenböck, H., Wurm, M., Netzband, M., Zwenzner, H., Roth, A., Rahman, A., & Dech, S. (2011). Flood risks in urbanized areas – multi-sensoral approaches using remotely sensed data for risk assessment. *Natural Hazards and Earth System Sciences, 11*, 431-444

Taubenböck, H., Wurm, M., Setiadi, N., Gebert, N., Roth, A., Strunz, G., Birkmann, J., & Dech, S. (2009). Integrating remote sensing and social science. In, *2009 Joint Urban Remote Sensing Event* (pp. 1-7)

Tewkesbury, A.P., Comber, A.J., Tate, N.J., Lamb, A., & Fisher, P.F. (2015). A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sensing of Environment, 160*, 1-14

Thompson, E.S., & de Beurs, K.M. (2018). Tracking the removal of buildings in rust belt cities with open-source geospatial data. *International Journal of Applied Earth Observation and Geoinformation, 73*, 471-481

Thonfeld, F., Feilhauer, H., Braun, M., & Menz, G. (2016). Robust Change Vector Analysis (RCVA) for multi-sensor very high resolution optical satellite data. *International Journal of Applied Earth Observation and Geoinformation, 50*, 131-140

Tian, J., Cui, S., & Reinartz, P. (2014). Building change detection based on satellite stereo imagery and digital surface models. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 406-417

Tigges, J., Lakes, T., & Hostert, P. (2013). Urban vegetation classification: Benefits of multitemporal RapidEye satellite data. *Remote Sensing of Environment, 136*, 66-75

Tobler, W., Deichmann, U., Gottsegen, J., & Maloy, K. (1997). World population in a grid of spherical quadrilaterals. *International Journal of Population Geography, 3*, 203-225

Tomás, L., Fonseca, L., Almeida, C., Leonardi, F., & Pereira, M. (2016). Urban population estimation based on residential buildings volume using IKONOS-2 images and lidar data. *International Journal of Remote Sensing, 37*, 1-28

Townshend, J., Justice, C., Li, W., Gurney, C., & McManus, J. (1991). Global land cover classification by remote sensing: present capabilities and future possibilities. *Remote Sensing of Environment, 35*, 243-255

Townshend, J.R.G., & Justice, C.O. (1988). Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations. *International Journal of Remote Sensing, 9*, 187-236

UN-Habitat (2016). *World Cities Report 2016: Urbanization and Development – Emerging Futures*. UN-Habitat

United Nations, Department of Economic and Social Affairs, Population Division (2014). *World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)*.

United Nations, Department of Economic and Social Affairs, Population Division (2017). *World Population Prospects: The 2017 Revision. Key Findings and Advance Tables*. Working Paper No. ESA/P/WP/248

Van Rijsbergen, C.J. (1979). *Information Retrieval*. Butterworths

Varshney, A., Arora, M.K., & Ghosh, J.K. (2012). Median change vector analysis algorithm for land-use land-cover change detection from remote-sensing data. *Remote Sensing Letters, 3*, 605-614

Veljanovski, T., Kanjir, U., Pehani, P., Oštir, K., & Kovačič, P. (2012). Object-based image analysis of VHR satellite imagery for population estimation in informal settlement Kibera-Nairobi, Kenya. *Remote Sensing-Applications*: InTech

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks, 11*, 586-600

Vitousek, P.M., Mooney, H.A., Lubchenco, J., & Melillo, J.M. (1997). Human Domination of Earth's Ecosystems. *Science, 277*, 494-499

Volpi, M., Tuia, D., Bovolo, F., Kanevski, M., & Bruzzone, L. (2013). Supervised change detection in VHR images using contextual information and support vector machines. *International Journal of Applied Earth Observation and Geoinformation, 20*, 77-85

Volpi, M., Tuia, D., Camps-Valls, G., & Kanevski, M. (2010). Unsupervised change detection by kernel clustering. In, *Proc. SPIE 7830, Image and Signal Processing for Remote Sensing XVI* (pp. 78300V-78300V-78308). Toulouse, France

Volpi, M., Tuia, D., Camps-Valls, G., & Kanevski, M. (2012). Unsupervised change detection with kernels. *IEEE Geoscience and Remote Sensing Letters, 9*, 1026-1030

Wang, B., Choi, S., Byun, Y., Lee, S., & Choi, J. (2015). Object-based change detection of very high resolution satellite imagery using the cross-sharpening of multitemporal data. *IEEE Geoscience and Remote Sensing Letters, 12*, 1151-1155

Wang, C., Wang, Y., Geng, Y., Wang, R., & Zhang, J. (2016a). Measuring regional sustainability with an integrated social-economic-natural approach: a case study of the Yellow River Delta region of China. *Journal of Cleaner Production, 114*, 189-198

Wang, J., Zhou, W., Qian, Y., Li, W., & Han, L. (2018a). Quantifying and characterizing the dynamics of urban greenspace at the patch level: A new approach using object-based image analysis. *Remote Sensing of Environment, 204*, 94-108

Wang, L., Li, C., Ying, Q., Cheng, X., Wang, X., Li, X., Hu, L., Liang, L., Yu, L., Huang, H., & Gong, P. (2012a). China's urban expansion from 1990 to 2010 determined with satellite remote sensing. *Chinese Science Bulletin, 57*, 2802-2812

Wang, M., Qi, S., & Zhang, X. (2012b). Wetland loss and degradation in the Yellow River Delta, Shandong Province of China. *Environmental Earth Sciences, 67*, 185-188

Wang, S., Tian, Y., Zhou, Y., Liu, W., & Lin, C. (2016b). Fine-Scale Population Estimation by 3D Reconstruction of Urban Residential Buildings. *Sensors, 16*, 1755-1781

Wang, X., Liu, S., Du, P., Liang, H., Xia, J., & Li, Y. (2018b). Object-Based Change Detection in Urban Areas from High Spatial Resolution Images Based on Multiple Features and Ensemble Learning. *Remote Sensing, 10*, 276-289

Warner, T.A., & Almutairi, A. (2009). Remote Sensing of Land Cover Change. In T. Warner, Nellis, M., & Foody, G. (Ed.), *The SAGE Handbook of Remote Sensing*. 55 City Road, London: SAGE Publications, Inc.

WBGU (2016). Humanity on the move: Unlocking the transformative power of cities. In. Berlin, Germany: German Advisory Council on Global Change

Webb, A.R. (2011). *Statistical pattern recognition*. (3rd ed.). Chichester, United Kingdom: John Wiley & Sons

Wen, Q., Zhang, Z., Shi, L., Zhao, X., Liu, F., Xu, J., Yi, L., Liu, B., Wang, X., Zuo, L., Hu, S., Li, N., & Li, M. (2016). Extraction of basic trends of urban expansion in China over past 40 years from satellite images. *Chinese Geographical Science, 26*, 129-142

Weng, Q., Quattrochi, D., & Gamba, P.E. (2018). *Urban Remote Sensing, Second Edition*. Boca Raton, FL, USA: Taylor & Francis

Weszka, J.S., Dyer, C.R., & Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-6*

Williams, D.P., Myers, V., & Silvious, M.S. (2009). Mine Classification With Imbalanced Data. *IEEE Geoscience and Remote Sensing Letters, 6*, 528-532

Wohlfart, C., Kuenzer, C., Chen, C., & Liu, G. (2016a). Social–ecological challenges in the Yellow River basin (China): a review. *Environmental Earth Sciences, 75*, 1066-1085

Wohlfart, C., Liu, G., Huang, C., & Kuenzer, C. (2016b). A River Basin over the Course of Time: Multi-Temporal Analyses of Land Surface Dynamics in the Yellow River Basin (China) Based on Medium Resolution Remote Sensing Data. *Remote Sensing, 8*, 186-210

Wong, D.F.K., Li, C.Y., & Song, H.X. (2007). Rural migrant workers in urban China: living a marginalised life. *International Journal of Social Welfare, 16*, 32-40

Woodworth, M.D., & Wallace, J.L. (2017). Seeing ghosts: parsing China's "ghost city" controversy. *Urban Geography, 38*, 1270-1281

World Bank (2013). Planning, Connecting, and Financing Cities - Now: Priorities for City Leaders. In. Washington, DC, USA: World Bank

World Bank, & Development Research Center of the State Council, the People's Republic of China (2014). Urban China: Toward Efficient, Inclusive, and Sustainable Urbanization. In. Washington, DC

Wu, C., & Murray, A.T. (2005). A cokriging method for estimating population density in urban areas. *Computers, Environment and Urban Systems, 29*, 558-579

Wu, F. (2015). *Planning for Growth: Urban and Regional Planning in China*. Taylor & Francis

Wu, G., & Chang, E.Y. (2005). KBA: kernel boundary alignment considering imbalanced data distribution. *IEEE Transactions on Knowledge and Data Engineering, 17*, 786-795

Wu, J., Wang, S., Zhang, Y., Zhang, A., & Xia, C. (2019). Urban landscape as a spatial representation of land rent: A quantitative analysis. *Computers, Environment and Urban Systems, 74*, 62-73

Wu, S.-S., Qiu, X., & Wang, L. (2005). Population Estimation Methods in GIS and Remote Sensing: A Review. *GIScience & Remote Sensing, 42*, 80-96

Wu, S.-S., Wang, L., & Qiu, X. (2008). Incorporating GIS Building Data and Census Housing Statistics for Sub-Block-Level Population Estimation. *The Professional Geographer, 60*, 121-135

Wurm, M., Schmitt, A., & Taubenböck, H. (2016). Building Types Classification Using Shape-Based Features and Linear Discriminant Functions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*, 1901-1912

Wurm, M., & Taubenböck, H. (2018). Detecting social groups from space – Assessment of remote sensing-based mapped morphological slums using income data. *Remote Sensing Letters, 9*, 41-50

Wurm, M., Taubenböck, H., Schardt, M., Esch, T., & Dech, S. (2011). Object-based image information fusion using multisensor earth observation data over urban areas. *International Journal of Image and Data Fusion, 2*, 121-147

Xian, G., Homer, C., & Fry, J. (2009). Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sensing of Environment, 113*, 1133-1147

Xiao, P., Wang, X., Feng, X., Zhang, X., & Yang, Y. (2014). Detecting China's Urban Expansion Over the Past Three Decades Using Nighttime Light Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7*, 4095-4106

Xiao, P., Yuan, M., Zhang, X., Feng, X., & Guo, Y. (2017). Cosegmentation for Object-Based Building Change Detection From High-Resolution Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing, 55*, 1587-1603

Xiao, P., Zhang, X., Wang, D., Yuan, M., Feng, X., & Kelly, M. (2016). Change detection of built-up land: A framework of combining pixel-based detection and object-based recognition. *ISPRS Journal of Photogrammetry and Remote Sensing, 119*, 402-414

Xie, Y., Gong, H., Lan, H., & Zeng, S. (2018). Examining shrinking city of Detroit in the context of socio-spatial inequalities. *Landscape and Urban Planning, 177*, 350-361

Xie, Y., Weng, A., & Weng, Q. (2015). Population Estimation of Urban Residential Communities Using Remotely Sensed Morphologic Data. *IEEE Geoscience and Remote Sensing Letters, 12*, 1111-1115

Xie, Y., & Weng, Q. (2016). Updating urban extents with nighttime light imagery by using an object-based thresholding method. *Remote Sensing of Environment, 187*, 1-13

Xu, R., & Wunsch, D.C. (2010). Clustering Algorithms in Biomedical Research: A Review. *Biomedical Engineering, IEEE Reviews in, 3*, 120-154

Xu, X., Peng, H., Xu, Q., Xiao, H., & Benoit, G. (2009). Land Changes and Conflicts Coordination in Coastal Urbanization: A Case Study of the Shandong Peninsula in China. *Coastal Management, 37*, 54-69

Xu, Z.X., Takeuchi, K., Ishidaira, H., & Zhang, X.W. (2002). Sustainability Analysis for Yellow River Water Resources Using the System Dynamics Approach. *Water Resources Management, 16*, 239-261

Yang, X. (2011). *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*. Wiley

Yang, X.t., Liu, H., & Gao, X. (2015). Land cover changed object detection in remote sensing data with medium spatial resolution. *International Journal of Applied Earth Observation and Geoinformation, 38*, 129-137

Ye, S., Chen, D., & Yu, J. (2016). A targeted change-detection procedure by combining change vector analysis and post-classification approach. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 115-124

Ye, X., Liu, M., Wang, J., Qin, Q., Ren, H., Wang, J., & Hui, J. (2017). Building-Based Damage Detection From Postquake Image Using Multiple-Feature Analysis. *IEEE Geoscience and Remote Sensing Letters, 14*, 499-503

Yeom, J., Jung, M., & Kim, Y. (2017). Detecting damaged building parts in earthquake-damaged areas using differential seeded region growing *International Journal of Remote Sensing, 38*, 985-1005

Yu, Q., Acheampong, M., Pu, R., Landry, S.M., Ji, W., & Dahigamuwa, T. (2018). Assessing effects of urban vegetation height on land surface temperature in the City of Tampa, Florida, USA. *International Journal of Applied Earth Observation and Geoinformation, 73*, 712-720

Zhang, B., Wang, R., Deng, Y., Ma, P., Lin, H., & Wang, J. (2019a). Mapping the Yellow River Delta land subsidence with multitemporal SAR interferometry by exploiting both persistent and distributed scatterers. *ISPRS Journal of Photogrammetry and Remote Sensing, 148*, 157-173

Zhang, C., Wei, S., Ji, S., & Lu, M. (2019b). Detecting Large-Scale Urban Land Cover Changes from Very High Resolution Remote Sensing Images Using CNN-Based Classification. *ISPRS International Journal of Geo-Information, 8*, 189

Zhang, H., Gong, M., Zhang, P., Su, L., & Shi, J. (2016). Feature-Level Change Detection Using Deep Representation and Feature Change Analysis for Multispectral Imagery. *IEEE Geoscience and Remote Sensing Letters, 13*, 1666-1670

Zhang, K.H., & Song, S. (2003). Rural–urban migration and urbanization in China: Evidence from time-series and cross-section analyses. *China Economic Review, 14*, 386-400

Zhang, Q., & Seto, K.C. (2011). Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data. *Remote Sensing of Environment, 115*, 2320-2329

Zhang, T., & Huang, X. (2018). Monitoring of Urban Impervious Surfaces Using Time Series of High-Resolution Remote Sensing Images in Rapidly Urbanized Areas: A Case Study of Shenzhen. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11*, 2692-2708

Zhang, X., Du, S., & Wang, Q. (2018a). Integrating bottom-up classification and top-down feedback for improving urban land-cover and functional-zone mapping. *Remote Sensing of Environment, 212*, 231-248

Zhang, X., Xiao, P., Feng, X., & Yuan, M. (2017). Separate segmentation of multi-temporal high-resolution remote sensing images for object-based change detection in urban area. *Remote Sensing of Environment, 201*, 243-255

Zhang, Y., Peng, D., & Huang, X. (2018b). Object-Based Change Detection for VHR Images Based on Multiscale Uncertainty Analysis. *IEEE Geoscience and Remote Sensing Letters, 15*, 13-17

Zhang, Y., & Sun, L. (2019). Spatial-temporal impacts of urban land use land cover on land surface temperature: Case studies of two Canadian urban areas. *International Journal of Applied Earth Observation and Geoinformation, 75*, 171-181

Zhang, Z., Liu, F., Zhao, X., Wang, X., Shi, L., Xu, J., Yu, S., Wen, Q., Zuo, L., Yi, L., Hu, S., & Liu, B. (2018c). Urban Expansion in China Based on Remote Sensing Technology: A Review. *Chinese Geographical Science, 28*, 727-743

Zheng, H., Wang, X., & Cao, S. (2014a). The land finance model jeopardizes China's sustainable development. *Habitat International, 44*, 130-136

Zheng, Q., Deng, J., Jiang, R., Wang, K., Xue, X., Lin, Y., Huang, Z., Shen, Z., Li, J., & Shahtahmassebi, A.R. (2017a). Monitoring and assessing “ghost cities” in Northeast China from the view of nighttime light remote sensing data. *Habitat International, 70*, 34-42

Zheng, Q., Zeng, Y., Deng, J., Wang, K., Jiang, R., & Ye, Z. (2017b). “Ghost cities” identification using multi-source remote sensing datasets: A case study in Yangtze River Delta. *Applied Geography, 80*, 112-121

Zheng, Y., Zhang, X., Hou, B., & Liu, G. (2014b). Using combined difference image and k-means clustering for SAR image change detection. *IEEE Geoscience and Remote Sensing Letters, 11*, 691-695

Zhou, W., Troy, A., & Grove, M. (2008). Object-based Land Cover Classification and Change Analysis in the Baltimore Metropolitan Area Using Multitemporal High Resolution Remote Sensing Data. *Sensors, 8*, 1613-1636

Zhu, J., Su, Y., Guo, Q., & Harmon, T.C. (2017). Unsupervised Object-Based Differencing for Land-Cover Change Detection. *Photogrammetric Engineering & Remote Sensing, 83*, 225-236

Zhu, Z., Zhou, Y., Seto, K.C., Stokes, E.C., Deng, C., Pickett, S.T.A., & Taubenböck, H. (2019). Understanding an urbanizing planet: Strategic directions for remote sensing. *Remote Sensing of Environment, 228*, 164-182

Eidesstattliche Erklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 42/2018 am 11.07.2018 angegebenen Hilfsmittel angefertigt habe.

Tobias Leichtle

München, den 01.07.2019